Experimentation and Approval Mechanisms

Andrew McClellan*

University of Chicago Booth School of Business

January 28, 2019

Abstract

We study the design of approval rules when costly experimentation must be delegated to an agent with misaligned preferences. When the agent has the option of ending experimentation, we show that, in contrast to standard stopping problems, the optimal approval rule must be history-dependent. We characterize optimal rule and show that the approval threshold moves downward over the course of experimentation. The threshold in force at any time depends on the history only via the minimum of previous beliefs and the current belief. We find that private information qualitatively may change the optimal mechanism: an agent can choose a fast-track option in the form of an initially depressed approval threshold. On expiry of the fast track the threshold jumps up, introducing more stringent standards. Our results provide a theoretical foundation for both the loosening of approval standards on longer experimentation paths and fast-track mechanisms.

Keywords: Dynamic mechanism design, experimentation, approval rules.

* Email: andrew.mcclellan@chicagobooth.edu.

I am deeply indebted to and would especially like to thank Debraj Ray and Ennio Stacchetti for their constant support, guidance and insightful suggestions throughout this project. I would also like to thank Simon Board, Dhruva Bhaskar, Eric Budish, Drew Fudenberg, Boyan Jovanovic, Emir Kamenica, Alessandro Lizzieri, Laurent Mathevet, Marco Ottaviani, David Pearce, Jacopo Perego, Ariel Rubinstein, Andrzej Skrzypacz and Can Urgun for helpful conversations and suggestions as well as seminar participants at NYU, Princeton, UPenn, USC, Minnesota, Chicago Booth, Yale SOM, UCL, UCSD, Columbia, UC Davis, Berkeley and Stanford.
1 Introduction

In many real world economic situations, decision makers face a trade-off between making a decision quickly and accurately. For example, when deciding whether or not to approve a drug, the FDA can mandate that companies conduct clinical trials to determine the efficacy and any side-effects the drug may have. In the approval process the FDA must trade off the need for haste (in order to alleviate the suffering of those currently afflicted) and the need for patience (so as to gather more information in order to prevent the use of harmful drugs by the public). A key element of this environment is that information is not only generated by nature but is controlled by an agent (the drug company) with incentives that are not aligned with that of the FDA: the drug companies which perform the clinical trials bear the cost of experimentation and may have different preferences on when to approve the drug. Thus the approval rule used by the FDA will determine how much experimentation the company is willing to perform. Additionally, the drug companies may possess private information which the FDA must elicit. For example, a company may spend a long time developing a drug prior to the start of a clinical trial and gets a private information about whether the drug is good or not. The misalignment of incentives will prevent straightforward elicitation: the company, which wants the drug to be approved more quickly, may have an incentive to exaggerate their optimism about the drug’s quality. This type of agency problem is present in many other settings, such as a venture capitalist deciding whether to invest in a start-up or a manager deciding whether or not to promote an employee. Understanding these agency considerations is important for determining the best approval rule.

In this paper, we revisit the canonical Wald hypothesis-testing problem with the new feature that approval and experimentation are controlled by separate players. We look at how a regulator can design stopping and decision rules (without monetary transfers) which incentivize an agent to perform experimentation and truthfully reveal any private information they have about the state of nature. The players have misaligned incentives in that the regulator prefers more experimentation before making a decision than the agent does. Thus the regulator has additional incentives to consider when designing his optimal stopping rule: in addition to the trade off between haste and discretion, the regulator must also consider how to continually incentivize the agent to continue experimentation. We study what the regulator’s optimal mechanism will be in the presence and absence of private information and under different levels of commitment and will show how splitting the control of approval and experimentation introduces rich and novel dynamics into the design of the optimal approval rule.
The contribution of this paper is several-fold: first, we look at a novel class of approval rules and prove their optimality among all approval rules. We find that these rules are history-dependent, giving us new insights into how agency considerations add a rich set of dynamics. They also generate interesting implications (e.g., longer experimentation is associated with more erroneous approval) and possess a number of desirable properties. Second, we study the effects of private information and find that it adds new qualitative features to the approval rules. Finally, we extend our model to encompass a large class of stopping games and show that the qualitative features of our optimal mechanism are quite general.

We start our analysis by investigating the case in which the agent has no private information (symmetric information), focusing on how the regulator provides incentives for the agent to experiment. A robust result from the optimal stopping literature with a single decision-maker is the optimality of stationary threshold strategies, in which the decision-maker stops whenever the state crosses a stationary, history-independent, threshold. If the incentives of the regulator and agent were aligned, such a rule would be optimal in our environment.

However, in many real life situations, the player producing information is different than the one using the information to make a decision and the production of information must be incentivized; if the clinical trials begin to go poorly and approval appears unlikely, the drug company may pull the plug on the clinical trial. In this setting, the regulator must consider interim participation constraints for the agent (which we call one-sided commitment). We show that under one-sided commitment, stationary threshold rules are no longer optimal: once the agent is about to quit, the regulator may have an incentive to change the approval rule so that the agent chooses to continue experimentation. However, there are many ways the regulator could change the approval rule to give the agent incentives for experimentation. Because the beliefs over the state of nature are changing over time, standard dynamic contracting methods become intractable. Given the richness of the set of approval rules, solving for the optimal rule can appear quite daunting. However, we find a tractable way to relax the problem and find the solution to be a novel optimal stopping rule which is history-dependent and non-stationary but still remarkably simple.

We show how the optimal mechanism can be written as a function only of the current belief and the minimum over the realized path of beliefs up to the current time. The approval rule consists of an approval threshold which moves downward sporadically; more specifically, the principal starts with a threshold which initially stays fixed as beliefs change. If beliefs descend low enough that the agent would optimally quit against the
current threshold if the current threshold were to remain fixed forever, then the principal begins to lower just enough to incentivize the agent to continue experimentation. When beliefs move higher than the current minimum, the threshold stays fixed (never increasing) and will only decrease when beliefs again reach a new low. This drift downwards of the approval threshold is bounded; if beliefs reach a lower fixed threshold, the regulator allows the agent to quit. Unlike in the case of a single decision-maker, the probability of Type I error is not constant over time. This mechanism possesses a number of attractive features; for example, we find that the optimal mechanism’s thresholds are independent of the initial beliefs, which would not be the case if we were to restrict the principal to only consider stationary threshold rules.

At a first pass, the optimal mechanism may seem to depend strongly on the assumption that the regulator can commit to the approval rule. To understand the role of regulator commitment, we examine what the regulator-optimal equilibrium is and find that our one-sided commitment mechanism is implementable without commitment. Whereas the previous literature has restricted attention to simple Markov equilibria which use only current beliefs as a state variable, this finding shows that, if we remove this particular restriction, we can implement the regulator-optimal equilibrium using only one additional state-variable (the minimum of beliefs up to the current time). We also show that every Pareto efficient equilibrium can be implemented using a similar mechanism to our own (with a different initial threshold). These equilibria are Pareto efficient after every history, giving them a robustness to renegotiation concerns.

Having explored the tension caused by the need to incentivize experimentation, we then introduce private information about the state and look at the new features this adds to the problem. When we give private information to the agent (i.e., a private signal about the state), we find that the optimal mechanism may take the form of a “fast-track” menu option. Low types select into a mechanism which is qualitatively similar to the case with no adverse-selection i.e., the approval threshold is monotonically decreasing when beliefs reach new lows. However, high types may be given a qualitatively different mechanism, a fast-track. In the fast track the agent is initially given a low approval threshold, but also faces a stationary “failure” threshold. If the failure threshold is reached, the project is not rejected but the approval threshold takes a discrete jump upward (they are thrown out of the fast-track); that is, they are allowed to continue to experiment but face a more stringent standard. This result shows how adverse selection creates a back-loading of costly distortions (raising the approval threshold) for the high type. By introducing a higher approval threshold, the regulator hurts both his and the agent’s payoffs. However, a deviating low type will view this distortion as more
likely. This allows the regulator to create separating contracts even without transfers and increase the probability of quicker approval for the high type.

Many of the results will go through for a large class of environments. In Section 6, we show that form of the optimal symmetric information mechanism, a threshold decreasing as beliefs reach new lows, holds in a much richer class of stopping games which allow for more general diffusion processes (not just a learning environment) as well as general payoff functions and outside options. This allows us to generalize our findings to apply to many more principal-agent problems and show that the dynamics we study are a feature of a wide class of such problems, which we illustrate with several examples.

In Section 2 we will discuss related literature and then introduce the model in Section 3. Section 4 will cover the optimal mechanism where there are no information asymmetries while Section 5 will derive the optimal mechanism when there are information asymmetries. Section 6 provides extensions and generalize the model to a wider class of diffusion processes and payoff functions.

2 Literature

The setting of our paper ties into a large literature on the problem of dynamic hypothesis testing. Wald (1947)’s seminal paper looked at the problem of sequential testing and began a rich literature in mathematics and statistics. Peskir and Shiryaev (2006) provide a textbook summary and history of the problem. Moscaroni and Smith (2001) also examine a similar framework but they look at the optimal policy in a large class of sampling strategies. Unlike our paper, this literature focuses on the problem of a single-decision maker. While some papers study the optimal stopping problem under constraints, the participation constraints our problem will impose are new and yield very different solutions.

A rich literature looking at the strategic forces in experimentation when many agents may produce information. Bolton and Harris (1999), Keller, Rady and Cripps (2005), Keller and Rady (2010, 2015), Strulovici (2010), Chan et al. (2015) and many others have analyzed the strategic interaction among experimenting agents. Typically, they focus on equilibrium experimentation levels and often find equilibrium strategies in cutoff rules. In our paper, we will endow one player (the regulator) with commitment power, which will drive the optimality of more complex stopping rules.

Closer to our own motivation, several papers have used this bandit framework to study the incentivization of experimentation. Garfagnini (2011) studies equilibrium levels of experimentation when a principal must delegate experimentation to an agent. Guo
(2016), one of the closest papers to our own, looks at a bandit problem in a principal-agent model when the agent possesses private information about the probability that the bandit is “good.” Like our model, Guo finds optimal mechanisms when monetary transfers are infeasible and the agent has private information about a payoff-relevant state of the world. Besides the technical differences between our settings, (Guo examines the optimal mechanism for eliciting information in a bandit model while we consider the optimal mechanism in a stopping problem, and in our model the misalignment between principal and agent preferences may be more severe), we consider the case in which the agent has the ability at any time to quit experimenting whereas in her model the principal controls experimentation throughout. Adding these interim constraints on the mechanism is what will add the history-dependence in our mechanism.

Another paper close to our own is Henry and Ottaviani (2018), who study a model of regulatory approval when learning takes place through a publicly observed Brownian motion. In their model, both the regulator and the agent possess a common prior about the state. They study equilibria of the approval process for different configurations of approval and experimentation authority. They find that varying the level of commitment (allowing players to potentially commit to a stationary threshold for approval or cessation of experimentation) and the possession of approval authority changes the expected amount of experimentation and study the social costs and benefits of different allocations of authority and commitment. In contrast, our main focus is on the design of optimal mechanisms and the effects of private information rather than equilibrium outcomes, which leads us to derive a new class of optimal mechanisms with a rich history dependence.

The incentivization of experimentation using monetary transfers from a moral hazard viewpoint has also been analyzed by Bergemann and Hege (1998,2005) and Horner and Samuelson (2013). Liu, Halac and Kartik (2016a, 2016b) also look at different ways of incentivizing experimentation, both in the framework of a contest and a contract. Our paper differs in that we are not allowing for monetary transfers and there is no moral hazard. Because our agent has intrinsic preferences over the action of the principal, the principal’s approval rule will be used to provide incentives in our model. Kruse and Strack (2015) look at an optimal stopping problem in a principal-agent framework in which the principal sets transfers in order to incentivize an agent to use particular stopping rules. They find that, under some conditions, transfers which only depend on the stopping decision implement cut-off rules and all cut-off rules are implementable by such transfers. Madsen (2016) also studies a principal-agent stopping problem with transfers in the case of the quickest detection problem.
The form of our optimal approval rule exhibits a kind of rigidity in how it decreases which is somewhat reminiscent of Holmstrom and Harris (1982). Unlike Holmstrom and Harris (1982) model, our rigidity arises from a dynamic contracting problem (rather than an equilibrium zero-profit condition). The movement downward of the threshold, towards the agent’s preferred level, is also reminiscent of backloading dynamics as seen in Ray (2002). Our principal, as beliefs drop, internalizes more of the agent’s utility moving the approval threshold towards the agent’s preferred level. Changing thresholds have been found in other optimal stopping papers, such as Fudenberg et al. (2018), who find a similar relationship between Type I error and the length of experimentation. However, their results are driven by the use of a different underlying state space rather than agency considerations as in our own paper, which we discuss in more detail in Section 4.

The study of the FDA approval process has also been studied theoretically and empirically Carpenter and Ting (2007) looks at a theoretical model of drug approval when the drug companies are better informed about the state for their drug. They study the resulting equilibria of a discrete time model. They find that the length of experimentation determines the comparative static on the effect of firm size on the amount of Type I and Type II errors. Carpenter (2004) also studied the effect of firm size on regulatory decisions. Frank et. al (2014) and Carpenter et. al (2008) look at the effects of regulatory changes at the FDA on the probability of Type I error.

3 Model

3.1 Environment

We begin by introducing our baseline model, which simplifies the analysis but as we will show in Section 6 can be generalized to a large degree. Following our motivating example, we study the interaction between a (female) regulator $R$ and a (male) agent $A$ in an infinite-horizon continuous-time model. Both players share a common discount rate $r > 0$. A project, which is up for approval, may be of two types: good ($\theta = H$) or bad ($\theta = L$). The regulator wants to approve only good projects. The benefit to approving a good project is $a_H$ and the loss to approving a bad project is $a_L$: 
We will assume that \( A \) pays a constant flow cost \( c_A \) until the game ends and \( R \) pays a flow cost of \( c_R \). For simplicity, we assume that \( c_A = c > 0 = c_R \) (none of the results will rely on \( c_R = 0 \), but this assumption makes the analysis simpler).

In general, the terminal payoffs of \( R \) and \( A \) might differ. For example, \( A \) might only care about the project being approved (if, for example, \( a^A_H = a^A_L = 1 \)). While we allow for general terminal payoffs (and extend the results further in Section 6), to simplify notation we take in the rest of the text \( a^R_H = a^A_H = 1 \) and \( a^R_L = -1 \leq a = a^A_L \leq 1 \) (we assume \( a \leq 1 \) so that \( A \) weakly prefers approval when \( \theta = H \) over approval when \( \theta = L \)). By changing \( a \), we vary the bias \( A \) has over terminal decisions from being aligned with \( R \) to always preferring approval. When \( a = -1 \) the only difference between \( A \) and \( R \)'s payoffs is that \( A \) bears the cost of experimentation. Taking \( a = -1 \) makes clear the difference in costs is key to the rich dynamics in the optimal mechanism.

Both players begin the game with a common-prior \( \pi_0 = \mathbb{P}(\theta = H) \). Over the course of the game, both players learn about the underlying state of nature \( \theta \) (which we call experimentation). Information about the state is revealed via a Brownian diffusion process with a state-dependent drift. Formally, while experimentation is ongoing, both players publicly observe

\[
X_t = \mu_{\theta t} + \sigma W_t
\]

where \( W = \{W_t, \mathcal{F}_t, 0 \leq t < \infty\} \) is a standard one-dimensional Brownian motion\(^1\) on the state space \((\Omega, \mathcal{F}, \mathbb{P})\) (where \( \omega \in \Omega \) is a sample point) and \( \mu_L = -\mu < 0 < \mu = \mu_H \). By observing \( X_t \), both players update beliefs about the state. After observing \( X_t \), a player’s posterior belief is given by Bayes rule as

\[
\pi_t = \frac{\pi_0 f^H_t(X_t)}{\pi_0 f^H_t(X_0) + (1 - \pi_0) f^L_t(X_t)},
\]

where \( f^\theta_t \) is the density of a normal distribution with mean \( \mu_{\theta t} \) and variance \( \sigma^2 t \). To simplify the belief updating procedure, we note that we can write the beliefs in terms of

\(^1\)Which implies that \( W_0 = 0 \) so \( X_0 = 0 \).
log-likelihoods\(^2\) - i.e.,

\[ Z_t = \log\left( \frac{\pi_t}{1 - \pi_t} \right). \]

Putting in our terms for \( \pi_t \), we have (after some algebra)

\[
Z_t = \log\left( \frac{\pi_0}{1 - \pi_0} \right) + \log\left( \frac{f_t^H(X_t)}{f_t^L(X_t)} \right) = Z_0 + \frac{\phi}{\sigma} X_t.
\]

where \( \phi = \frac{2\mu}{\sigma} \), which is called the signal-to-noise ratio, describes how informative the signals are. This transformation of the belief process is useful because both \( X_t \) and the initial \( Z_0 \) enter linearly into the current \( Z_t \). Since beliefs and the evidence level are isomorphic, we will use them interchangeably in the following sections.

As discussed before, in some situation it is reasonable that the agent has more information before the public news process begins. We model such situations by allowing the agent to receive an initial private signal to time \( t = 0 \) which is private information.

**Definition 1.** The model has **symmetric information** if the agent’s private signal is uninformative about the state. The model has **asymmetric information** if the initial signal is informative.

Note that in a model with asymmetric information, all the private information of \( A \) is realized at \( t = 0 \); that is, all information observed by \( A \) after \( t = 0 \) is also observed by \( R \).

We define \( \mathcal{F}_t^X = \sigma((X_s, Y_0) : 0 \leq s \leq t) \) (where \( Y_0 \sim U[0, 1] \) time 0 is used simply to allow for randomization) to be the augmented natural filtration and assume it satisfies standard restrictions (see Karatzas and Shreve (1991)). A history \( h_t = \omega|_{[0,t]} \) is the realization of a path of \( X_t \) (from time 0 to \( t \)) and \( Y_0 \).

We choose to model the news process as Brownian motion for both tractability and its similarity to real-world applications. In our motivating example, if \( X_t \) corresponds to patient’s health during a clinical trial, then the choice of Brownian motion reflects the gradual nature of learning and the noisiness of health outcomes. Even when administered good drugs, a patient’s health will still sometimes decline. However, the drift of a patient’s health should be positive for good drugs (i.e., \( \mu_H > 0 \)). The use of Brownian motion ties into a rich statistics literature on the design of adaptive clinical trials and

\(^2\)We subsequently abuse notation by referring to \( Z_t \) as beliefs.
hypothesis testing. We assume that the signal is publicly observable to both R and A. This assumption is satisfied in many situations. For example, the FDA can require companies to publicly register and continuously report the outcome of the trial. Assuming the news process is public allows us to avoid the situation in which R and A’s beliefs diverge over time, which would make the model intractable.

3.2 Mechanism

Our goal is to understand how R can optimally design approval standards to both incentivize the agent to perform experimentation and to elicit the private information (if there is any). We will assume that transfers are infeasible as is the case in many real-world settings. Formally, we allow R to design a stopping mechanism, which consists of a stopping time and a decision rule to approve or reject conditional upon stopping:

Definition 2. A stopping mechanism is a pair $\langle \tau, d_{\tau} \rangle \in \mathcal{T} \times \mathcal{D}$, where $\mathcal{T}$ is the set of $\mathcal{F}_t^X$-measurable stopping rules and $\mathcal{D}$ is the set of $\mathcal{F}_t^X$-measurable decision rules taking values in $\{0, 1\}$.

When discussing our stopping mechanisms, it will be useful to discuss how the mechanism behaves after a particular history $h_t$.

Definition 3. For stopping mechanism $\langle \tau, d_{\tau} \rangle$ and history $h_t$, the continuation mechanism at $h_t$ is $\langle \tau[h_t], d_{\tau[h_t]} \rangle$ and is defined for each $\omega$ with history $h_t$ by $\tau[h_t](\zeta_t(\omega)) = \tau(\omega) - t$ and $d_{\tau[h_t]}(\zeta_t(\omega)) = d_{\tau}(\omega)$, where $\zeta_t : \Omega \to \Omega$ is the shift operator defined such that $X_t(\omega) = X_0(\zeta_t(\omega))$.

The agency considerations in the model will impose constraints on the mechanisms which are allowed, which will impose constraints on which problems we consider admissible.

Definition 4. Let $\Delta_C \subseteq \mathcal{T} \times \mathcal{D}$ and define the constrained problem $CP$ to be

$$\sup_{\langle \tau, d_{\tau} \rangle \in \Delta_C} \mathbb{E}[e^{-r\tau}g(X_{\tau}, d_{\tau})|X_0].$$

We say that $\langle \tau, d_{\tau} \rangle$ is admissible with respect to $CP$ if $\langle \tau, d_{\tau} \rangle \in \Delta_C$.

For most of the paper we will endow R with perfect commitment power, allowing us to focus on direct-revelation stopping mechanisms in the asymmetric information model, and we assume that the decision to approve or reject is irrevocable.\(^3\) The utility of R for a particular mechanism $\langle \tau, d_{\tau} \rangle$ is given by

\(^3\)This irrevocability assumption is without loss if we allow experimentation to stopped and restarted and the agent only pays his flow cost while experimentation is ongoing.
\[ J(\tau, d_{\tau}, Z_0) = \mathbb{E}[e^{-r\tau}(\pi_{\tau} - (1 - \pi_{\tau}))d_{\tau} | Z_0] = \mathbb{E}[e^{-r\tau}\frac{e^{Z_{\tau}} - 1}{1 + e^{Z_{\tau}}}d_{\tau} | Z_0] \]

and the utility of \( A \) is given by

\[ V(\tau, d_{\tau}, Z_0) = \mathbb{E}[e^{-r\tau}(\pi_{\tau} + a(1 - \pi_{\tau}))d_{\tau} - \int_0^\tau e^{-rt}cdt | Z_0] = \mathbb{E}[e^{-r\tau}(d_{\tau}\frac{e^{Z_{\tau}} + a - c}{1 + e^{Z_{\tau}} - r}) | Z_0] - \frac{c}{r} \]

Before moving on the general analysis, we first define some notation that will be useful in the following analysis. We begin with a salient subclass of mechanisms, in which the mechanism is characterized by a pair of thresholds: the regulator approves if her beliefs ever reach \( B \) and rejects in her beliefs ever reach \( b \). We will refer to \( B \) as the static approval threshold and \( b \) as the static rejection threshold.

**Definition 5.** A static threshold mechanism is a pair \((b, B) \in \mathbb{R}^2 \) such that \( b \leq Z_0 \leq B \). \( \tau = \inf\{t : Z_t \notin (b, B)\} \) and \( d_{\tau} = \mathbb{I}(Z_{\tau} \geq B) \).

This focal class of stopping mechanisms are tractable, easily implemented and have the useful property that they allow us to calculate the expected utility for \( R \) and \( A \) in closed form. To express these utilities, we must know the expected discounted probability that a threshold is reached. The formula\(^4\) for the discounted probability of reaching \( B \) before \( b \) when the \( \theta = H \) is given by

\[ \Psi(B, b, Z) := \mathbb{E}[e^{-r\tau}d_{\tau} | \theta = H, Z_0 = Z] = \frac{e^{-R_1(Z - b)} - e^{-R_2(Z - b)}}{e^{-R_1(B - b)} - e^{-R_2(B - b)}}, \]

and the discounted probability that the beliefs cross \( b \) before ever crossing \( B \) if \( \theta = H \) is

\[ \psi(B, b, Z) := \mathbb{E}[e^{-r\tau}(1 - d_{\tau}) | \theta = H] = \frac{e^{R_2(B - Z)} - e^{R_1(B - Z)}}{e^{R_2(B - b)} - e^{R_1(B - b)}}, \]

where \( R_1 = \frac{1}{2}(1 - \sqrt{1 + \frac{2r\sigma^2}{\mu^2}}) \) and \( R_2 = \frac{1}{2}(1 + \sqrt{1 + \frac{2r\sigma^2}{\mu^2}}) \).

Doing a bit of algebra (see Henry and Ottaviani (2018)) allows us to show that the discounted probability that \( B \) is crossed before \( b \) if \( \theta = L \) is

\[ \Psi(B, b, Z)e^{Z - B} \]

and the the discounted probability that \( b \) is crossed before \( B \) if \( \theta = L \) is

\[ \psi(B, b, Z)e^{Z - b} \]

\(^4\)See Stokey (2009).
This allows us to rewrite the utility of $R, A$ when $(\tau, d_{\tau})$ takes a static threshold form as

$$\tilde{J}(B, b, Z_0) := J(\tau \geq (B) \land \tau \leq (b), 1(Z_\tau = B), Z_0) = \frac{e^{Z_0}}{1 + e^{Z_0}} \Psi(B, b, Z_0)(1 - e^{-B})$$

$$\tilde{V}(B, b, Z_0) := V(\tau \geq (B) \land \tau \leq (b), 1(Z_\tau = B), Z_0),$$

$$= -\frac{c}{r} + \frac{e^{Z_0}}{1 + e^{Z_0}} \left( \Psi(B, b, Z_0)(1 + \frac{c}{r} + (a + \frac{c}{r})e^{-B}) + \frac{c}{r}\psi(B, b, Z_0)(1 + e^{-b}) \right).$$

We will generally drop the dependence of $\Psi, \psi$ on $B, b, Z$ when the choice of $B, b, Z$ is clear. To simplify notation, we will also use $\Psi_b := \frac{\partial \Psi}{\partial b}$ and $\Psi_B := \frac{\partial \Psi}{\partial B}$ (with similar notation for the derivatives of $\psi$).

### 4 Symmetric Information

Our set-up features two-types of tensions: that the agent must be incentivized to experiment and that the agent may have private information. In order to understand how each of these tensions affect the regulator’s choice of a mechanism, it is useful to understand how they each work separately before adding them both to the model. We will therefore begin our analysis by shutting down the private information component and focus on the design of optimal mechanisms under symmetric information-i.e., both $A$ and $R$ share the same prior when the news process begins. Studying the symmetric information case will be useful both for finding the optimal mechanism with asymmetric information and of independent interest. Our model extends the canonical hypothesis-testing model, which is well-studied in single decision-maker problems, into a mechanism-design framework in which the decision maker faces an additional trade-off in that she incentivize $A$ to continue to experiment. Additionally, we examine how optimal mechanisms change depending on the level of commitment of $R, A$. Exploring this dimension yields new dynamics in the optimal mechanism.

#### 4.1 One-Sided Commitment

If $R$ could control the length of experimentation, it is straightforward to show that she would employ a stationary threshold strategy. This finding turns out to be robust to even mild forms of the agency problem. In the situation where $R$ can force $A$ to commit to the length of experimentation (what we call two-sided commitment), we show in Appendix G that $R$ would optimally employ a stationary threshold. However, in many settings assuming that $A$ can be forced to continue experimenting is unrealistic.
If, over the course of the trial, the company becomes pessimistic that the drug will ever be approved, a drug company may decide to cut their losses and end the trial early. While the FDA can commit to approval standards, the ability to compel the company to continue running a clinical trial is beyond the scope of the agency’s authority. Thus, A must be incentivized to continue experimentation and not take his outside option even after \( t = 0 \). We call this environment one of one-sided commitment.

We also allow for A, once R has approved to immediately quit and take his outside option of zero rather than the payoff for approval. Under the case of commitment, since R would never find it optimal to direct A to stop experimenting and restart later, it is without loss to for now take quitting experimentation to be irreversible (although we will relax this when we drop commitment later).

Since A has the ability at any time to take an outside option, we must ensure that A’s continuation payoff is weakly positive at all \( t \) and histories \( h_t \) until R ends experimentation. The mechanism must then satisfy a dynamic version of the usual participation constraint.

**Definition 6.** A mechanism \( (\tau, d_\tau) \) satisfies the **dynamic participation constraint** if after any history \( h_t \) the expected continuation to A from \( (\tau[h_t], d_\tau[h_t]) \) is non-negative:

\[
\forall h_t, \quad E[e^{-r(\tau[h_t])}(d_\tau[h_t]\frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_t, h_t] - \frac{c}{r} \geq 0.
\]

Intuitively, another way of stating the idea behind the dynamic participation constraint is to say that A never finds it strictly optimal to quit. Suppose that the agent chooses to quit before R ends experimentation—i.e., A chooses a quitting rule \( \tau' \in \mathbb{T} \) by which he takes his outside option of 0 at time \( \tau' \). This strategy will give A an expected utility of

\[
E[e^{-r(\tau \land \tau')}](d_{\tau'}1(\tau \leq \tau')\frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) - \frac{c}{r}.
\]

Following this idea and restricting R to choose from mechanisms which incentivize A to not quit prematurely, we define R’s problem as

\[
[SM] : \sup_{(\tau, d_\tau)} E[e^{-r(\tau \land \tau')}d_{\tau}\frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}}|Z_0]
\]

subject to

\[
DP : \sup_{\tau' \in \mathbb{T}} E[e^{-r(\tau \land \tau')}d_{\tau'}1(\tau \leq \tau')\frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_0] \leq E[e^{-r(\tau_{\tau'}d_{\tau'}\frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_0].
\]
**DP** implies that for any quitting rule that $A$ might use, the payoff to potentially quitting early is weakly less than letting $R$ decide when to end experimentation. Specifying that the optimal mechanism must not let $A$ quit before $R$ approves or rejects is without loss: If a mechanism allows $A$ to quit after a history $h_t$, then we could specify another mechanism in which $R$ rejects at the same moment that $A$ quits. This will not change any incentives for $A$ to quit earlier than time $t$ and hence the expected payoff to $R$ from the two mechanisms will be the same. As we formally prove in Lemma 7 in the Appendix, the constraint **DP** is essentially a rewritten version of the definition of dynamic participation constraints from an ex-ante perspective and so we can focus on solving **SM**.\footnote{With a slight abuse of notation, we will refer to **DP** as dynamic participation constraints for the rest of the paper.}

Given the previous literature, it seems natural to conjecture the optimality of static threshold mechanisms. Surprisingly, we find that the conjecture fails, as we illustrate with a simple example below to show how threshold rules can be improved upon. Simply put, whenever $R$ rejects in a static threshold mechanism, she would be better off lowering the threshold (“cutting the $A$ some slack”) in order to incentivize $A$ to continue experimenting. By changing the threshold after some histories, $R$ is better able to fine tune the incentives for experimentation to $A$.

Suppose that $a > 0$ and $R$ is using a static approval threshold of $B_1 > 0$. Note that $R$ receives positive utility from approving at belief $Z_t$ if and only if $Z_t \geq 0$. We will refer to $Z_t = 0$ as $R$’s myopic cutoff point—i.e., the belief at which she would approve if she were myopic. If $B_1 < 0$, then the static threshold mechanism would only approve at beliefs which give $R$ negative utility and, therefore, $R$ would be better off rejecting immediately. Since $R$ would always benefit from continued experimentation at all beliefs below this approval threshold, he will never reject the project before the agent would decide to quit. Let $b^*_Z$ be the value at which $A$ will choose to quit experimenting when $R$ uses a static threshold of $B$ and the current beliefs are $Z$. In general, this choice will depend on the approval threshold $B$. It is straightforward to show that the argmax over $b$ of $\tilde{V}(B, b, Z_0)$ is independent of $Z_0$ and thus the optimal quitting threshold is only a function of $B$. More formally, we define $b^*_Z$ as

\[
 b^*_Z(B) := \arg\max_b \tilde{V}(B, b, Z_0),
\]

In order to satisfy **DP** it must be that $R$ rejects when $Z_t = b^*_Z(B)$.

Let Mechanism 1 by a static-threshold mechanism $(B_1, b^*_Z(B_1))$. The expected payoff to $R$ from this mechanism will be
However, upon reaching $b^*_Z(B_1)$, $R$ rejects the project and takes her outside option. At this point she would be better off if she could convince $A$ to keep experimenting. Now consider Mechanism 2, in which $R$ uses an approval threshold of $B_1 > 0$ until either $Z_t = B_1$ or $Z_t = b^*_Z(B_1)$. If $Z_t$ reaches $b^*_Z(B_1)$ first, then, instead of rejecting, $R$ lowers the approval threshold to $B_2 \in (0, B_1)$. $A$ will now continue experimenting until the evidence reaches $b^*_Z(B_2) < b^*_Z(B_1)$, since the lowering of the approval threshold strictly incentivizes $A$ to keep experimenting. Under this new policy, the expected payoff to $R$ from Mechanism 2 is

$$
\Psi(B_1, b^*_Z(B_1), Z_0) \frac{e^{Z_0}(1 - e^{-B_1})}{1 + e^{Z_0}}.
$$

This is multiplied by the discounted probability that beliefs hit $b^*_Z(B_1)$ before $B_1$

$$
\psi(B_1, b^*_Z(B_1), Z_0) \frac{e^{Z_0}(1 + e^{-b^*_Z(B_1)})}{1 + e^{Z_0}}.
$$

Note that $\Psi(B_2, b^*_Z(B_2), b^*_Z(B_1)) \frac{e^{b^*_Z(B_1)(1 - e^{-B_2})}}{1 + e^{Z_1}}$ is strictly positive (since $e^{-B_2} < 1$). Therefore Mechanism 2 yields a higher payoff for $R$. Since the choice of $B_1$ was arbitrary, we can see that any static-threshold mechanisms are not optimal.

Once we have moved out of the realm of threshold rules, conjecturing the form that the optimal policy will take is difficult. Because the space of stopping rules is large, it is...
not clear if there is salient class of mechanisms that the optimal policy will lie in or if the optimal policy is feasible to derive. The key difficulty comes from the sup over \( \tau' \in \mathbb{T} \) in the DP constraint. Unlike standard mechanism design, where the agent can deviate by misreporting along some one-dimensional interval, the DP constraint allows the agent to deviate across an infinite-dimensional class. Overcoming this difficulty is the main challenge of this section. By finding a tractable relaxed version of SM, we are able to solve for the optimal mechanism and show that it possesses a relatively simple structure. Roughly speaking, the optimal strategy will be a continuous version of strategy above, where \( R \) decreases the threshold whenever \( A \) is about to quit, but keeps fixed as beliefs move back toward the threshold.

In the interest of keeping notation consistent throughout the following sections, we will describe the mechanism in terms of \( X_t \) rather than \( Z_t \).\(^6\) We define an equivalent version of \( b^*_Z \) for the process measured in terms of \( X_t \) and a value \( B^*_A \) which would be the optimal choice of approval threshold (in \( X_t \)-space) for \( A \) if he could choose the static approval threshold:

\[
\begin{align*}
b^*(B) &:= b^*(B; Z_0) := [b_Z(Z_0 + \phi B) - Z_0] \frac{\sigma}{\phi}, \\
B^*_A &:= B^*_A(Z_0) := [\text{argmax}_{B} \tilde{V}(B, b^*_Z(B), Z_0) - Z_0] \frac{\sigma}{\phi}.
\end{align*}
\]

We find that the optimal mechanism turns out to depend on the realized path of \( X_t \) only through the current minimum of the evidence path \( M^X_t := \min \{X_s : s \in [0, t]\} \). We can describe the optimal mechanism as consisting of two regimes: A initial stationary regime which uses static approval threshold \( B^1 \) which lasts until \( X_t \) reaches \( B^1 \) or \( b^*(B^1) \) and a second incentivization regime where the stopping rule is given by the first time \( X_t \) crosses \( B(M^X_t) \) which decreases as \( M^X_t \) decreases in order to incentivize \( A \) to keep experimenting when beliefs get too low. \( R \) decreases the current threshold in order to incentivize \( A \) to keep experimenting; the decrease is both gradual (just enough to keep \( A \) from quitting) and permanent (the threshold never increases, only decreases). The optimal mechanism continues the sporadic decrease of the approval threshold \( B(M^X_t) \) from \( R \)'s preferred level to the maximum of \( A \)'s preferred level and \( R \)'s myopic threshold and has the notable property that once the threshold has decreased, it will never rise again.

Let us define \( B(X) \) as the lowest static approval threshold \( B \) above \( A \)'s preferred threshold \( B^*_A \) such that \( A \) would choose to quit at evidence level \( X \) when \( R \) is using a

\(^6\)This will be useful when we introduce asymmetric information so that we don’t have to describe the mechanism in terms of the beliefs of both \( A \) and \( R \).
static threshold mechanism with approval threshold $B$-i.e.,

$$B(X) := \min\{B > B_A^* : b^*(B) = X\},$$

Formally, the resulting optimal mechanism is stated in the following theorem.

**Theorem 1.** The optimal stopping mechanism is given by the stopping rule which approves the first time $X_t$ goes above the function $B(M_t^X)$ and rejects the first time it falls below a fixed threshold $b$. Formally, the $\tau = \inf\{t : X_t \notin (b, B(M_t^X))\}$ and $d_\tau = 1(X_\tau \geq B(M_t^X))$ where $B(M_t^X)$ is defined as

$$B(M_t^X) = \begin{cases} B^1 & M_t^X \in [b^*(B^1), 0] \\ B(M_t^X) & M_t^X \in [b^*(B^1)). \end{cases}$$

and $b = b^*(B_A^*) \lor b^*(-\frac{\sigma}{\alpha}Z_0)$. When translating the mechanism into belief space, the optimal threshold $B_{Z}(M_t^X)$ is independent of $Z_0$.

The main feature of this mechanism is the non-stationarity of the approval threshold. Since both $R$ and $A$ would choose stationary thresholds if they controlled both approval and experimentation, we see that the dynamics are driven by the agency problem caused by the preference misalignment. This type of history dependence in the approval rule is, to our knowledge, new in the context of this widely used stopping problem and shows how strategic interactions can lead to a rich set of dynamics in the design of approval rules.

The approval threshold decreases until either it reaches $A$’s preferred level $B_A^*$ (which happens when $M_t^X = b^*(B_A^*)$) or it reaches $R$’s myopic threshold (which happens when $M_t^X = b^*(-\frac{\sigma}{\alpha}Z_0)$). $A$ can never be incentivized to experiment at beliefs below $b^*(B_A^*)$: in his first best, $A$ would be quitting at $Z < b^*(B_A^*)$, a payoff he can replicate even when he doesn’t have control of the approval threshold by quitting immediately. Moreover, $R$ will

---

The assumption of a binary state is important for the optimality of a threshold which is stationary in the *expected payoff of approval* (here $Z_t$) in the single-decision maker benchmark; a continuous state space would lead to a richer approval rule (e.g., Fudenberg et al. 2018) who study the case a normal prior and show that the mean belief (i.e., expected payoff from approval) at which the decision maker approves decreases over time). However, in such models, the approval rule is stationary in the beliefs over the full state space; that is, for the same beliefs over the state space, the optimal approval rule will be the same independent of the history to get there (e.g., in the normal specification, the sufficient state-space is that of the mean and variance of posterior beliefs, where variance decreases deterministically over time. In the binary state space, the beliefs at any given time are given by $Z_t$). This independence is not present in our model: different histories leading to the same belief may use different continuation approval rules.
Figure 2: The graph above corresponds to the changing approval threshold for a particular realization of $X_t$. The upper dashed line corresponds to the current approval threshold. This approval threshold will stay at the same level until $X_t$ crosses the current minimum of the process, which is given by the bottom dotted line.

never choose to lower the approval threshold below her myopic threshold $-\frac{\sigma}{\phi} Z_0$ (since doing so would only guarantee her a negative payoff and she would be better off letting the agent quit). Thus experimentation is not extended indefinitely but ends whenever the approval threshold reaches either the agent optimal level or $R$’s myopic cutoff.

Another notable feature about the optimal mechanism is that the approval threshold, when written in terms of $Z_t$ rather than $X_t$ does not depend on the initial belief of $Z_0$, only $Z_t$ and $M^{Z}_t$ (proven formally in Lemma 8 in the Appendix). This property is common in single-decision maker problems. However, it is absent if we were to restrict attention to a choice over static threshold rules (see Henry and Ottaviani (2018)). Thus the moving away from just static approval thresholds restores a common feature of the single decision-maker framework to the solution when considering the agency problem.

The rest of this section will be devoted to sketching out the ideas of the proof. The tractability afforded by continuous time has led to a growing literature in mechanism and contract design. Our approach differs from the standard continuous-time approach (e.g., Sannikov (2007), where transfers are feasible, and Fong (2007), where transfers are not feasible) where agent-continuation payoffs are formulated as a state variable in an HJB equation. Because they are learning about the underlying state, using the HJB approach in our model would require carrying both a state variable of agent continuation and
Figure 3: The dashed line gives the approval threshold as a function of $M^X$ and the solid straight line marks the 45 degree line where $X_t = M^X$. The dashed line is initially constant in $M^X$ during the stationary regime while it decreases in $M^X$ for the incentivization regime. The lines coming up from the 45 degree line illustrate a sample path of $X_t$. In this case, the process reached the approval threshold when $X_t = 0.7$ and $M^X = -1.4$.

current beliefs about the state; finding the solution to the HJB equation would require solving a difficult partial differential equation and is impractical for analyzing the optimal stopping rule. Instead, we use a different approach by finding a relaxed problem over which Lagrangian techniques work well. This method allows to more easily derive the qualitative features of the optimal mechanism with which we are able to explicitly pin down the form of the optimal mechanism.

A key difficulty lies in the fact that when checking $DP$, we must consider all possible quitting rules $\tau'$ which $A$ might use. For an arbitrary stopping rule $(\tau, d_\tau)$, we might try to solve for the optimal $\tau'$ which $A$ would use. However, even for simple time-dependent stopping rules, solving for $\tau'$ is difficult and cannot be calculated in closed-form. Given the richness of the set of available $(\tau, d_\tau)$, which may be history-dependent, solving for $\tau'$ is infeasible. Given the dimensionality of the space of quitting rules, it is not immediately clear how to do this.

In order to simplify the problem, we need to make some conjectures on when we expect $A$ to find it most profitable to quit experimenting. In Appendix G, we show that
if $R$ can force $A$ to commit to the amount of experimentation, then the optimal policy is a static threshold mechanism. This shows us that $R$ finds it optimal to “smooth” incentives to $A$ by lowering the approval threshold but not by violating stationary. As long as $R$ is using a stationary threshold, then $A$’s best response quitting rule would be a stationary threshold. This intuition for smoothing of the approval threshold by $R$ leads us to conjecture a particular class of $\tau'$ deviations will be binding, those which we call threshold quitting rules:

**Definition 7.** A uses a **threshold quitting rule at** $X_i$ if he quits at time $\tau \leq (X_i)$ (we henceforth drop the $\leq$ in the subscript for notational convenience).

The payoffs to $A$ of quitting early are equal to those of rejection. Therefore, $A$ evaluates the $(\tau, d_\tau)$ when following quitting rule $\tau(X_i)$ as equivalent to the mechanism $(\tau \wedge \tau(X_i), d(X_i))$ where we define $d(X_i) := d_\tau \mathbb{1}(\tau \leq \tau(X_i))$. We then study a relaxed problem in which we restrict attention to a finite number of such quitting rules. Let $\mathcal{T}_N = \{X_i\}_{i=0}^N$ such that $X_0 = 0$ and $X_{i+1} = X_i + \frac{X}{N}$ for some $X \in \mathbb{R}_-$ so that the solution to two-sided commitment problem starting at $Z_0 = \frac{-\phi}{\sigma} X$ would be immediate rejection. The restriction to a finite grid of points is done for technical reasons and we will look at the limit as this grid becomes arbitrarily fine. Formally we define our relaxed problem $RSM_N$ by replacing the $DP$ constraint in $SM$ with a set of relaxed $RDP$ constraints (one for each $X_i \in \mathcal{T}_N$):

$$RDP(X_i) : \mathbb{E}[e^{-r(\tau \wedge \tau(X_i))}(d_\tau(X_i) \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_0] \leq \mathbb{E}[e^{-r\tau}(d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_0].$$

We note that because we have dropped a number of constraints (i.e., all non-threshold quitting rules), the solution to $RSM_N$ will provide an upper bound on the value to $R$ of the full problem $SM$.

We should emphasize that it is not obvious that dropping non-threshold constraints is without loss. For many stopping policies $R$ could use, the best response of $A$ will not be to use a threshold policy. For example, if $R$ were to wait until date $T$ and approve if and only if $X_T > B$, then the optimal quitting rule $A$ would use would in fact not be a threshold policy but would be a time-dependent curve $\tau' = \inf \{t : X_t = f(t)\}$. Since we allow for arbitrarily complex history-dependent stopping rules, the quitting rule which is $A$’s best response to an arbitrary $\tau$ may also be a complex history-dependent quitting rule. We should also note that we are not restricting the solution of $RSM_N$ to be a threshold policy. Instead, we are only checking that $A$ has no incentive to deviate to a threshold quitting rule rather than obediently following $R$’s proposed mechanism.
We can now transform our primal problem $RSM_N$ into its corresponding dual problem by constructing an associated Lagrangian with Lagrange multipliers $\{\lambda(X_i)\}_{i=0}^N \in \mathbb{R}^{N+1}$:

$$
\mathcal{L} = \sup_{(\tau,d)} E[e^{-r\tau}d \frac{e^{Z_\tau}}{1+e^{Z_\tau}}|Z_0] \\
+ \sum_{i=0}^N \lambda(X_i)(E[e^{-r(\tau\wedge\tau(X_i))}(d_\tau(X_i) \frac{e^{Z_\tau} + a}{1+e^{Z_\tau} + \frac{c}{r}}|Z_0] - E[e^{-r\tau}(d_\tau \frac{e^{Z_\tau} + a}{1+e^{Z_\tau} + \frac{c}{r}}|Z_0])].
$$

For an appropriate choice of $\{\lambda(X_i)\}_{i=0}^N$, the solution to the associated Lagrangian will solve the primal problem $RSM_N$ and will appropriate have complementary slackness conditions.\(^8\)

This dual version of the problem drastically simplifies the analysis since it allows us to study what is effectively a single decision-maker problem. And while the selection of appropriate multipliers $\{\lambda(X_i)\}_{i=0}^N$ is difficult, the qualitative properties that we will derive from the analysis of the Lagrangian for arbitrary multipliers will allow us to pin down the form of the optimal solution.

Let $X^1$ be the first binding threshold. We will decompose the problem into the time before $X^1$ has been reached and the time after the first quitting rule has been reached (i.e., $\tau(X^1)$). If $A$ is truly indifferent between quitting and continuing at $\tau(X^1)$, then his continuation value at $\tau(X^1)$ should be zero. We denote the value of the optimal mechanism for $R$ (satisfying the relaxed $RDIC$ constraints) starting at $X_t$ which delivers an expected value of 0 to $A$ by $H_N(X_t)$ (this value doesn’t depend on the history before time $t$, only the starting value $X_t$).

With the problem in an unconstrained form, we can use techniques from the single decision-maker stopping problem to find the optimal policy which solves the dual problem. The following lemma allows us to establish the optimality of a “local” static-threshold rule: the approval threshold stays constant until the first binding constraint $X^1$ is reached. Moreover, we show that $A$ will truly be indifferent between quitting and continuing at $\tau(X^1)$.

**Lemma 1.** The solution to $RSM_N$ is a stationary threshold approval policy until $\tau(X^1)$. The continuation mechanism at $\tau(X^1)$ is the solution to $H_N(X^1)$.

Lemma 1 establishes the optimality of the initial stationary regime in $RSM_N$. This result doesn’t contradict our earlier result that static thresholds are non-optimal: as

\(^8\)Strong duality follows from Dokuchaev (1997) and Balzer and Janben (2002), formally stated in Lemma 5.
we show below, the mechanism in the second regime will not turn out to be a static threshold mechanism. A key thing to note is that the value of $H_N(X^1)$ is independent of the history up until time $\tau(X^1)$. What happens after $\tau(X^1)$ is completely bundled into the value $H_N(X^1)$ and therefore doesn’t affect the choice of $R$ before $\tau(X^1)$ except through the value of $H_N(X^1)$. This also implies that whenever $X^1$ is reached, $A$ is indifferent between quitting and continuing to experiment. This property comes from the particular form of the optimal stopping rule and is not true of all stopping rules that $R$ could use. For example, if $R$ was using a deterministic stopping rule $\tau = T$ for some $T \in \mathbb{R}_+$, then it could be that $A$’s $RDP(X^1)$ constraint was binding in expectation at $t = 0$, but, when $X^1$ is first reached, $A$ could have a strictly positive or negative continuation value. The stationarity of the optimal stopping rule prior to $\tau(X^1)$ is key for proving the indifference of $A$ at $\tau(X^1)$.

We then solve for the optimal mechanism in the second regime (i.e. that which solves the problem which deliver $H_N(X^1)$). The basic argument mirrors that of Lemma 1. The difference is that when we are trying to solve for the optimal mechanism which deliver $R$ payoff $H_N(X^1)$ (and $A$ a payoff of zero), we have added a promise keeping constraint to the $RDP$ constraints. This additional constraint can be incorporated in a Lagrangian similar to the $RDP$ constraints and, by repeated application of the arguments used in Lemma 1, we can show the optimal mechanism to be a decreasing sequence of thresholds in $M_t^X$.

The proof establishes that for each $N$, the stopping mechanism depends on the history only through $(X_t, M_t^X)$. This property, which we will see repeatedly in the next session as well, is crucial for establishing that our relaxed problem is a solution to the full problem. While the mechanism is history-dependent, the mechanism can be determined with only two state-variables, making the mechanism tractable for calculation and implementation.

The most notable feature about the optimal mechanism is that the approval threshold is changing with $M_t^X$. For lower $M_t^X$, the lower approval threshold increases the probability of Type I error, in contrast to static threshold mechanisms in which the probability of error conditional upon approval is constant. Our model predicts that the analyst will predict a higher probability of Type I error (i.e., approving a bad project) for projects which have taken a long time to be approved when compared to projects which were approved quickly. In many contexts this fits a natural intuition. For example, if an assistant professor receives tenure very quickly, he is more likely to be judged to be of high quality than if he took a long time to receive tenure.
4.2 No Commitment

At first glance, the optimal mechanism under one-sided commitment seems to require a great deal of commitment: $R$ agrees to permanently lower the approval threshold, even though she would be better off raising in back to its initial level if beliefs drift back up. It is natural to think then that without commitment, she would be tempted to raise the threshold. However, this would destroy the promised incentives from when $A$ was about to quit, thereby unravelling the optimal mechanism. We might then naturally wonder how much $R$ loses if she cannot commit to the optimal mechanism. To answer this question, we need to think about the exact details of the model without commitment. More specifically, we need to know the precise sequence of events when $A$ stops experimenting: as we will show, these details are crucial for determining the equilibrium outcome. We describe several different set-ups below:

- (I): $A$ can irrevocably quit experimenting at any time $t$ and $R$ cannot approve after $A$ has quit.
- (II): $A$ can irrevocably quit experimenting at any time $t$ and $R$ can approve at any time after the agent quits.
- (III): $A$ can temporarily stop experimenting at any time. When $A$ is not experimenting, $A$ pays no flow cost and $R$ can approve at any time.

\textbf{Figure 4}: The likelihood of Type I error at a given approval at time $t$ is increasing.
Set-ups (I) and (II) have been studied in Kolb (2016) and Henry and Ottaviani (2018) and, if we restrict the Markov Perfect Equilibrium (MPE) using $Z_t$ as a state variable, have a static threshold structure. Let us therefore focus on the case of set-up (III).

Because we are moving out the realm of commitment, we must be clear about what we mean by an equilibrium. Formally, a strategy for the agent is a mapping from histories (or filtration) into the choice of whether or not to experiment: $a : \mathcal{F}_t \to \{0, 1\}$. Both agents observe $X_t$ which solves to stochastic differential equation $dX_t = a_t(\mu_\theta dt + \sigma dW_t)$. $R$’s strategy is given as before by a stopping time and decision rule $(\tau, d_\tau)$.

Definition 8. A pair $(a^*, (\tau^*, d^*_\tau))$ is an equilibrium if for every history $h_t$, the continuation actions $a^*[h_t]$ and $(\tau^*[h_t], d^*_\tau[h_t])$ satisfy

- A maximization: $a^*[h_t] = \arg\max E[e^{-r(\tau^*[h_t]-t)}d^*_\tau[h_t] - \int_{\tau^*[h_t]}^{T^*[h_t]} e^{-r(s-t)}a_s c|h_t]$.
- $R$ maximization: $(\tau^*, d^*_\tau) = \arg\max E[e^{-r(\tau^*[h_t]-t)}\int_{\tau^*[h_t]}^{T^*[h_t]} e^{-r(s-t)}a_s c|h_t]$.
- Belief Updating: $dZ_t = \frac{\phi}{\sigma} dX_t$.

One natural justification for the restriction to MPE is that they are “simple” enough to be implementable in real world situations and minimize history dependence. However, as we will show below, this restriction is with loss if we are looking for the $R$-optimal equilibrium. It is natural to worry that when deriving this optimal equilibrium, we may need to use complex history dependence in the strategies. Surprisingly, we show that a complex structure is not necessary the optimal equilibrium: if we only slightly expand the state space to be $(X_t, M^X_t)$, then the optimal mechanism under one-sided commitment can be implemented as an MPE without commitment.

Proposition 1. Under set-up (III), the optimal mechanism under one-sided commitment can be implemented as an equilibrium.

The intuition behind the proof is quite simple. Suppose that $A$ expects $R$ to follow the mechanism as outlined in Theorem 1 and $R$ expects $A$ to continue experimenting until $R$ approves or $X_t = b^*(B^*_A) \lor b^*(-\frac{\phi}{\sigma}Z_0)$. Then $R$ has no incentive to approve early (if approving early were a profitable deviation, then she could implement it in the mechanism with commitment and still satisfy all DP constraints) or reject early (since rejection is always suboptimal). Moreover, $A$ has no incentive to cease experimenting early since his continuation value is always weakly positive (we specify off-path beliefs that if $A$ quits experimenting before called to, $R$ believes that $A$ will begin experimentation again

$^9(\tau, d_\tau)$ is taken to be measurable with respect to $a_t$. 

24
immediately, giving $A$ no incentive to delay). Moreover, if $A$ ceases experimentation at $X_t = B(M_t^X)$ until $R$ approves, $R$ also has no incentive to delay approval. The key to the proof is that neither player ever has a strictly negative continuation value and thus has no incentive to deviate from the prescribed equilibrium actions.

Proposition 1 tells us that our solution to the no-commitment case is the preferred equilibrium for $R$. We can also ask what other payoffs are generated by various equilibria. Our techniques from the case with commitment can be useful in generating the Pareto frontier of the equilibrium set. More formally, consider the problem of solving the principal’s problem so that it respects $DP$ constraints but also adds a promise keeping constraint to deliver at least expected utility $W$ to $A$:

\[ PK : \mathbb{E}[e^{-r\tau}(d_\tau e^{Z_\tau} + \frac{a}{1 + e^{Z_\tau}} + \frac{c}{r})] \geq W \]

It is straightforward from Theorem 4 to show that the solution to the above problem will take the form $\tau = \inf \{ t : X_t \notin (b, B(M_t^X)) \}$ (where $b = b^*(B_A^\ast) \vee b^*(-\frac{a}{c}Z_0)$) and $d_\tau = 1(X_\tau \geq B(M_t^X))$ where $B(M_t^X)$ is defined as

\[ B(M_t^X) = \begin{cases} 
B & M_t^X \in [b^*(B), 0] \\
B(M_t^X) & M_t^X \in [b^*(B_A^\ast) \vee b^*(-\frac{a}{c}Z_0), b^*(B)]. 
\end{cases} \]

for some $B \in [B_A^\ast \vee -\frac{a}{c}Z_0, B^1]$. Thus, by adjusting the initial threshold $B$ (which is equivalent to adjusting $W$), we can map out the Pareto frontier of the equilibrium set. Additionally, these mechanisms are Pareto optimal after every history, giving them a high degree of robustness to renegotiation.

5 Asymmetric Information

The previous section shows us how splitting control of approval and control of experimentation introduces a rich set of dynamics to the optimal approval rule. Having a better understanding of how this agency problem drives the dynamics, we can now build in our other natural tension, that of asymmetric information. In many principal-agent situations the agent may have private information about the state. For example, if $A$ is a start-up and $R$ a venture capitalist deciding when to invest, it is likely that $A$ is more informed than $R$ about the start-up’s profitability. In the case of drug companies, the company may have acquired information about the drug during the R&D phase or during animal or foreign clinical trials which are not directly observable by the FDA. This
information is valuable to $R$, as it will allow them to shorten experimentation time and make more informed decisions. In this section we study how to elicit this private information and look at what distortions it introduces. As we will see, the optimal mechanisms are qualitatively different when compared to the symmetric information case. When $A$ reports a higher prior on $\theta$, the optimal mechanism may entail giving him a fast-track to approval, where he is given a low initial approval threshold with the caveat that he may be thrown out of the fast-track if the outcomes of the trial go poorly and thereafter face a more stringent burden of proof for approval.

Intuitively, $R$ would like to set a lower approval threshold when $A$ reports a higher prior. However, since $A$ prefers a lower approval threshold than $R$, this introduces incentives for types with lower priors to misreport. In order to restore incentives, $R$ will seek to “punish” outcomes which lower types find more likely. Due to the different beliefs of $R$ when evaluating the mechanism for a high prior and a low prior $A$ evaluating the same mechanism, $R$ can “back-load” punishments in such a way as to reduce low type incentives while minimizing ex-ante distortions for $R$. In the optimal mechanism, this punishment comes in the form of being thrown out of the fast-track: $R$, knowing that $A$ has a lower prior, views the chances that the fast-track is revoked (which entails an inefficient increase in the approval threshold from $R$’s perspective) to be lower than $A$ does when $A$ has a lower prior and has deviated.

We model asymmetric information by allowing for the agent’s starting belief $\pi_A$ to take on a binary realization $\pi_A \in \{\pi_\ell, \pi_h\}$ where $\pi_\ell < \pi_h$. Translating into log-likelihood space, we will call the case when $A$ begins with prior $Z_0 = Z_\ell = \log(\frac{\pi_\ell}{1-\pi_\ell})$ the low type of $A$ (who we refer to as $\ell$) and when $A$ begins with the prior $Z_0 = Z_h = \log(\frac{\pi_h}{1-\pi_h})$ as the high type of $A$ (who we refer to as $h$). We let $P(Z_i)$ be the ex-ante probability of type $Z_i$.\footnote{The binary assumption is done for tractability. As we will see, it is already difficult to determine which incentive constraints will bind with just two types.}

We now redefine a stopping mechanism to account for the need to elicit the private information of the agent. By the Revelation Principle, we focus on direct mechanisms in which $A$ reports his type to $R$. This will result in $R$ offering a menu of stopping mechanisms from which $A$ can choose by reporting his type.\footnote{We are moving back to allowing $R$ to commit to the mechanism; unlike the symmetric information case, commitment will be necessary for implementing the solution.}

**Definition 9.** A stopping mechanism is a menu $\{(\tau^i, d^i_\tau)\}_{i=h,\ell}$ such that $(\tau^i, d^i_\tau) \in T \times D$ and $R$ implements $(\tau^i, d^i_\tau)$ when she receives message $m = i$.

When we are considering the effects of $A$ misreporting his type, the beliefs of $A$ and
$R$ will be different. Note that because initial beliefs enter linearly into $Z_t$, after any realization of $X_t$, the log-likelihood beliefs of $A$ and $R$ (when $A$ misreports his type) will be different by a constant $\Delta_Z := Z_h - Z_\ell$. This difference in the beliefs will be important as it implies that $R$ evaluates the probability of a mechanism’s outcomes differently than a deviating $A$ will.

5.1 One-Sided Commitment

We will need to define a dynamic version of incentive compatibility in a similar manner as we defined the dynamic participation constraints. Incentive constraints for type $i$’s value of reporting to be type $k$ must take into account that $i$ also considers the value of a deviation where he may choose to quit early.

Definition 10. A stopping mechanism under one-sided commitment is **dynamically incentive compatible** if for all $i, k$,

$$\sup_{\tau' \in \mathbb{T}} \mathbb{E}[e^{-r(\tau^k \land \tau')}(d^k_\tau \mathbb{1}(\tau < \tau') e^{Z_\tau} + \frac{a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_i] \leq \mathbb{E}[e^{-r\tau^i}(d^i_\tau e^{Z_\tau} + \frac{a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_i].$$

We include the $\sup$ over $\tau' \in \mathbb{T}$ in the incentive constraint to convey the fact that $i$ is comparing correctly declaring his type to be $i$ to the payoff he could get from reporting to be type $k$ and potentially quitting early. These types of double deviations (misreporting one’s type and quitting early) will play an important role in determining the optimal mechanism. The full problem for $R$ is then

$$[AM]: \sup_{(\tau', d^i_\tau) = i, h} \sum_{i=\ell, h} \mathbb{E}[e^{-r(\tau^i \land \tau')} (d^i_\tau e^{Z_\tau} + \frac{a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_i] \cdot \mathbb{P}(Z_i)$$

subject to $\forall i = h, \ell$ and $k \neq i$

$$DP(Z_i): \sup_{\tau'} \mathbb{E}[e^{-r(\tau^i \land \tau')}(d^i_\tau \mathbb{1}(\tau^i \leq \tau') e^{Z_\tau} + \frac{a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_i] \leq \mathbb{E}[e^{-r\tau^i}(d^i_\tau e^{Z_\tau} + \frac{a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_i]$$

$$DIC(Z_i, Z_k): \sup_{\tau'} \mathbb{E}[e^{-r(\tau^k \land \tau')} (d^k_\tau \mathbb{1}(\tau^k \leq \tau') e^{Z_\tau} + \frac{a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_i] \leq \mathbb{E}[e^{-r\tau^i}(d^i_\tau e^{Z_\tau} + \frac{a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_i].$$

We will proceed by analyzing the problem type-by-type. Unlike many mechanism design problems, there is no clear answer to which are the relevant constraints. In fact, different combinations of binding $DIC$ constraints will bind depending on the specific values of $Z_h, Z_\ell$. We begin by looking at what the optimal mechanism for $h$ is when $DIC(Z_\ell, Z_h)$ is binding and $DIC(Z_h, Z_\ell)$ is slack. This conjecture is natural when $Z_h$
is high; in this case, $R$ would like to give $h$ a lower approval than $\ell$; however, this may create an incentive for $\ell$ to misreport his type to get quicker approval and save on experimentation costs. We should note that this intuition is not always correct; for example, if $Z_\ell$ is low enough, for some parameter values both $R$ and $\ell$ would always prefer that $\ell$ quit immediately. Such a case would immediately revert back to the symmetric information model with initial belief $Z_0 = Z_h$.

Let $V_\ell$ be the utility that $\ell$ gets from truthfully declaring his type (this is determined by $R$ through his choice of $\ell$’s mechanism, but for now we can treat it as fixed.). Then our (relaxed) problem $AM^h$ of determining the optimal high type mechanism is given by $SM$ with the addition of the DIC constraint:

$$DIC(Z_\ell, Z_h, V_\ell) : \sup_{\tau} \mathbb{E}[e^{-r(\tau^h + \tau')}(d_\tau \mathbf{1}(\tau \leq \tau') \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_\ell] \leq \frac{c}{r} + V_\ell.$$  

Intuitively, we should expect that $h$’s DP constraints will not be binding as long $\ell$ has not found it strictly optimal to quit experimenting: since $h$ has a higher belief about the state being good, he will ascribe a higher probability to approval than $\ell$ would. The lower belief on $\ell$’s part means that for low enough $X_t$, the symmetric information mechanism for $h$ will induce $\ell$ to quit immediately. Let $(\tau_{SM}^{\ell h}, d_{SM}^{\tau, h})$ be the stopping rule for the symmetric mechanism given $h$ and define $b_{SM}^h$ to be the highest $X_t$ such that $(\tau_{SM}^{\ell h}, d_{SM}^{\tau, h})$ starting at $X_t$ would induce $\ell$ to quit immediately—i.e.,

$$b_{SM}^h := \max\{X_t \text{ such that} \sup_{\tau'} \mathbb{E}[e^{-r(\tau_{SM}^{h}\wedge\tau')}](d_{\tau'}^{\ell h} \mathbf{1}(\tau_{SM}^{h} < \tau') \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|X_t, M_t^X = X_t, Z_\tau = Z_\ell + \frac{\phi}{\sigma} X_\tau] = \frac{c}{r}\}.$$  

It is straightforward to see that such a $b_{SM}^h$ exists; for example, if $X_t$ is such that the optimal approval threshold for $h$ under symmetric information is given by $B_{h}(M_t^X)$, then $\ell$ will find it optimal to quit immediately since $B_{h}(M_t^X) > B_{j}(M_t^X)$.

Because $A$ cares about the state when $a < 0$, it is possible in some cases that $R$ could dissuade $\ell$ from misreporting to be $h$ by giving him too low of an approval threshold (i.e., threatening him with early approval). We find such threats unnatural and unappealing and they aren’t robust to situations in which $A$ could keep experimenting after approval. In order to rule out such threat of early approval, we will assume that at $b_{SM}^h$, $\ell$ would still prefer a lower approval threshold. More specifically, we assume that there is only one approval threshold which would leave $\ell$ indifferent between continuing to experiment and quitting.
Assumption 1. For each $b > b^S_M$ and $X > b$, there exists a unique $B_Z$ such that
$$
\tilde{V}(B_Z, Z_{\ell} + \frac{a}{\alpha} b, Z_{\ell} + \frac{a}{\alpha} X) = 0.
$$

If we consider payoffs such that $A$ prefers approval regardless of the state (i.e., $a > 0$) then such an assumption clearly holds. Alternatively, we can also rule out such threats with a low approval threshold without this assumption on $b^S_M$ as long as $A$ is allowed to continue experimenting (and delay approval or rejection until $A$ desires to do so) after $R$ approves (that is $A$’s payoff to approval is at $X_t$ is $\max_{\bar{b}} \tilde{V}(B, b, Z_{\ell} + \frac{a}{\alpha} X)$ rather than $\frac{e^{Z_t} a}{1 + e^{Z_t}}$). In such a situation in which $A$ effectively controls both experimentation and approval once $R$ has signed off on approval, $A$ can always get at least his preferred thresholds $B^*_A b^*_A$, making threats of “too-early” approval by $R$ non-credible and will ensure that $A$’s utility is weakly decreasing in the approval threshold and therefore there is a unique approval threshold which leaves the agent indifferent with taking his outside option.\footnote{In the case of drug companies, this means that the regulator can not forbid the companies from running additional trails after approval but before putting the drug on the market, which is a reasonable assumption.}

We can incorporate this into our analysis formally by replacing $A$’s terminal utility $\frac{e^{Z_t} a}{1 + e^{Z_t}}$ with $\max \{\frac{e^{Z_t} a}{1 + e^{Z_t}}, \tilde{V}(B^*_A, b^*_A, Z_t)\}$ and $R$’s utility with $\max \{\frac{e^{Z_t} a}{1 + e^{Z_t}}, \tilde{J}(B^*_A, b^*_A, Z_t)\}$.

Upon $\tau(b^S_M)$, using the symmetric mechanism for $h$ is optimal: the optimal mechanism for $AM^h$ will promise $\ell$ and $h$ types who have continued to this point some (weakly positive) continuation utilities and will respect all $DP$ constraints for $h$. Since the symmetric mechanism for $h$ does not have such promised continuation values and does respect all $DP$ constraints, it’s value will yield an upper bound on the original mechanism at time $\tau(b^S_M)$.

We solve this problem in a manner very similar to the symmetric information case and begin by defining a relaxed problem in which we drop all $DP$ constraints and all but a finite number of $DIC$ constraints and restrict the process to stop at $X_t = b^S_M$ with continuation value to $R$ equal to his payoff from the symmetric mechanism for $h$ at $X_t$ (which we will denote $SM^h(X_t)$). Formally, our (even more) relaxed problem can be written as

$$
[RAM^h_N]: \sup_{(r,d)} \mathbb{E}[e^{-r(\tau \land \tau(b^S_M))} (d_r(b^S_M) e^{Z_{\tau}\ell} - \frac{1}{1 + e^{Z_{\tau}}} \mathbb{I}(\tau \leq \tau(b^S_M)) + \mathbb{I}(\tau > \tau(b^S_M)) S M^h(b^S_M)| Z_h)]
$$

subject to $\forall X_i \in T_N \cup \{b^S_M\}$

$$
RDIC_e(X_i) : \mathbb{E}[e^{-r(\tau \land \tau(X_i) \land \tau(b^S_M))} (d(X_i) e^{Z_{\tau}\ell} + a \frac{1}{1 + e^{Z_{\tau}}} \mathbb{I}(\tau \leq \tau(b^S_M)) + \frac{c}{r})| Z_e] \leq V_t + \frac{c}{r}.
$$

The intuition behind solving this problem is very similar to that of solving the relaxed problem in $SM$. The main difference is the addition of the $DIC$ constraints for $\ell$. We
show that at the first binding constraint \( X^1 \), \( R \) must be using a stopping rule which induces \( \ell \) to weakly prefer to quit. There are many ways in which the optimal mechanism could induce \( \ell \) to quit while still providing incentives for \( h \) to experiment (e.g., there exists thresholds such that, due \( \ell \)'s lower belief, \( \ell \) would quit while \( h \) would continue to experiment). If the first binding constraint is \( b^{SM} \), then the symmetric information mechanism will induce \( \ell \) to quit.

It is important to note that while the continuation mechanism at \( \tau(X^1) \) must induce \( \ell \) to quit, the payoff relevant beliefs for \( R \) are those of \( h \). By inducing \( \ell \) to quit, the mechanism may be setting a stricter approval policy that \( R \) would like to given that the true beliefs are \( h \). If this is the case, then \( R \) would like to “relax” the overly stringent approval policy over time while making sure to do in such a way as to not violate the earlier incentives for \( \ell \) to quit. We verify that this intuition is correct: \( R \) loosens the threshold by decreasing the threshold as \( M_t^{X} \); as \( \ell \)'s beliefs get lower, a lower approval threshold is needed to ensure that \( \ell \) found it optimal to quit at \( X^1 \). First, though, we define some notation:

\[
b^\ast_i(B) := b^\ast(B; Z_i) \\
B^\ast_i(A) := B^\ast(A; Z_i) \\
B_i(X) := \min\{B > B^\ast_i(A) : b^\ast_i(B) = X\}.
\]

The function \( b^\ast_i(B) \) gives the threshold at which type \( i \) would quit when facing a static approval threshold is \( B \). \( B_i(X) \) is the approval threshold which would induce type \( i \) to quit at evidence level \( X \). With these definition in hand, we can formally state the mechanism for \( h \):

**Proposition 2.** When \( \text{DIC}(Z_h, Z_\ell) \) is slack, the solution to \( AM_h^h \), for some \( (b_1^h, B_1^h) \), the stopping mechanism \( \tau = \inf\{t : X_t \not\in (b, B^h(M_t^{X}))\} \) and \( d_\tau = 1(X_\tau = B^h(M_t^{X})) \) where

\[
B^h(M_t^{X}) = \begin{cases} 
B_1^h & \text{if } M_t^{X} \in [b_1^h, 0] \\
B_\ell(M_t^{X}) & \text{if } M_t^{X} \in [b^{SM}, b_1^h] \\
B^h_\ell(M_t^{X}) & \text{if } M_t^{X} < b^{SM}.
\end{cases}
\]

where \( B^{SM}_h(M_t^{X}) \) is symmetric information threshold function for \( h \) and \( b = b^\ast(B^\ast_h) \lor b^\ast_\ell(-\frac{\sigma}{\phi} Z_h) \).

The approval threshold \( B^h(M_t^{X}) \) is effectively pinned down by \( (b_1^h, B_1^h) \). When \( b_1^h > b_\ell^\ast(B_\ell^h) \), then \( B_\ell(b_1^h) > B_1^h \), implying that the approval threshold takes a jump when
$M^X_\ell$ crosses $b^1_h$. This feature distinguishes it from our symmetric information mechanism and, as we will see, the mechanism for $\ell$. The reason behind this jump follows from our discussion at the outset of the section 5.1: $b^1_h$ acts a punishment or failure threshold which, if reached, moves into a punishment phase in order to lower the incentives of $\ell$ to misreport his type.

This second stage of the optimal mechanism (i.e., once $b^1_h$ has been reached) has a continuous and monotonically decreasing (in $M^X_\ell$) threshold, which consists of two regimes. After exiting the initial stationary regime, the mechanism may enter a “punishment” regime in which the approval threshold jumps up. The fact that $B^h(M^X_\ell) = B^\ell(M^X_\ell)$ ensures that $\ell$ does indeed want to quit, but uses the minimal increase in the approval threshold necessary. $R$ would like to decrease the approval threshold as quickly as possible, but must satisfy $PK$ constraints which ensure that $\ell$ did indeed find it optimal to quit at $\tau(b^1)$. We show that the optimal means to do this is by using the threshold $B^\ell(M^X_\ell)$. At $\tau(b^{SM})$ the optimal mechanism moves into the symmetric mechanism regime. In this sense, the distortions introduced by asymmetric information dissipate over time since conditional on no approval before $M^X_\ell = b^{SM}$, eventually $R$ will be able to implement her best mechanism under one-sided commitment absent any distortions from private information, which implies that there is no distortion from private information at the end of experimentation.

The $DIC(Z_\ell, Z_h)$ constraint does not cause $R$ to reject at a higher belief that would be optimal in the absence of incentive constraints. This comes about because the $\ell$ always has more pessimistic beliefs than $h$ and thus there is always a way to deliver incentives for $\ell$ to quit that don’t involve rejection. This implies that the placement of distortions (relative to the symmetric information mechanism) are non-monotonic in $M^X_\ell$: they increase upon entering the punishment phase but then are gradually reduced over time until they are eliminated upon entering the regime using the symmetric information, in which $R$ uses the best mechanism he could get absent private information.

Having found what the optimal mechanism is for $h$, we must also solve for the optimal mechanism for $\ell$. Suppose that the mechanism must deliver utility $V_\ell$ to the low type correctly declaring his type. Then the mechanism design problem is given by the same problem as in the symmetric case but with the addition of a promise-keeping constraint ensuring that $\ell$’s expected utility is at least $V_\ell$:

$$PK(V_\ell) : \mathbb{E}[e^{-rt}(d_t e^{Z_\ell} + a + c_r Z_\ell)] \geq V_\ell.$$  

We should expect that the optimal mechanism for $\ell$ is qualitatively the same as the
The optimal mechanism for \( \ell \) when \( DIC(Z_h, Z_\ell) \) is slack is given by, for some \( b_r \), the stopping rule \( \tau = \inf \{ t : X_t \notin (b_r, B_\ell(M_t^X)) \} \) and \( d_\tau = \mathbb{1}(X_\tau = B_\ell(M_\tau^X)) \) which is defined as

\[
B_\ell(M_t^X) = \begin{cases} 
B_\ell^1, & M_t^X \in [b_\ell^*(B_\ell^1), 0], \\
B_\ell^2(M_t^X), & M_t^X \in [b_r, b_\ell^*(B_\ell^1)]. 
\end{cases}
\]

where \( B_\ell^1 \) is such that \( B_\ell^1 \) is less than it would be in the symmetric information case.

Proposition 3 leaves open the possibility that \( R \) rejects at \( b_r > b_\ell^*(B_\ell^1) \)-i.e., the second part of \( B_\ell(M_t^X) \) is never reached. Whenever \( A \) and \( R \)'s payoffs are aligned over terminal payoffs (i.e., \( a = -1 \)), then \( R \) will never reject early and \( b_r = b_\ell^*(B_{A,\ell}^1) \). However, this may not be true when \( a = 1 \). Although there are no \( DIC \) constraints to consider, \( R \), when designing \( \ell \)'s mechanism, does consider how increasing the utility of \( \ell \) when truthfully reporting weakens the incentives for \( \ell \) to misreport himself to be \( h \). Therefore, \( R \) will decrease the approval threshold lower than she would prefer; if this decrease is large enough, \( R \) may prefer to not enter the incentivization regime and instead reject at \( b_r \). However, it is straightforward to show that if \( P(Z_\ell) \) is large enough and \( Z_\ell > b_\ell^*(\frac{a}{\sigma} Z_\ell) \), then we will have \( b_\ell^1 = b_\ell^*(B_\ell^1) \). This will also always be the case when \( Z_\ell > 0 \).

Having derived the form of each mechanism for each type separately, we can now formally state the optimal mechanism.

Theorem 2. When \( DIC(Z_\ell, Z_h) \) is binding and \( DIC(Z_h, Z_\ell) \) is slack, the optimal mechanism is given the mechanisms of Propositions 2 and 3. If the optimal mechanism for \( h \) is not equal to the symmetric information mechanism and \( \ell \) is not forced to quit experimenting early (i.e., \( b_\ell^1 = b_\ell^*(B_\ell^1) \)), then \( B_\ell^1 \leq B_h^1 \) and \( B_\ell^1 < B_h^1 \) implies \( b_\ell^1 < b_h^1 \).

We refer to the mechanism given to \( h \) when \( B_h^1 < B_\ell^1 \), as a fast-track mechanism. We can think of \( h \) as being offered a two stage trial: the first trial (a fast-track) is given a low approval threshold, but also a “failure” threshold \( b_h^1 \). If the failure threshold is reached first, then the trial is declared a failure and the agent is thrown out of the fast-track. However, instead of rejecting, \( R \) allows \( h \) to immediately continue experimenting, only now \( h \) is given a more stringent approval threshold.

This fast-track mechanism illustrates the trade offs that must be made under one-sided commitment: in order to grant \( h \) a lower approval threshold, \( R \) must deter deviations by \( \ell \) by increasing the failure threshold. This lower approval threshold is more
Figure 5: If the initial regime thresholds were such that $B^1_\ell > B^1_h$ and $b^1_\ell > b^1_h$, then $\ell$ report himself to be $h$ and quit early, effectively giving himself a lower approval threshold. likely to be reached by $h$ than $\ell$, which allows $R$ to profitably back load distortions in the “failure” threshold.

We now illustrate the behind why the optimal mechanism must take this nested form when $b_r < b^*_r(B^1_\ell)$. For $\ell$, when declaring himself to be $\ell$ or $h$, his utility is completely determined by the initial static thresholds. Suppose that $Z_\ell > 0$ and $a > 0$; then we know that $b^*_\ell = b^*_r(B^1_\ell)$-i.e., $R$ keeps the initial threshold fixed until the point at which $\ell$ is first indifferent between ceasing and continuing experimentation. We note that if $B^1_h > B^1_\ell$, then we cannot have $DIC(Z_h, Z_\ell)$ binding. The reason for this shown in Figure 5: since the static approval threshold is higher (which strictly reduces utility to $\ell$), $\ell$ must gain from experimenting longer on the low end of beliefs when claiming to be $h$. But since $\ell$, when truthfully declaring his type, is allowed to experiment up until the point at which he would choose to quit, there is nothing to be gained (relative to truthfully declaring his type) for $\ell$ from claiming to be $h$. If $B^1_h < B^1_\ell$, then it must be that $b_h > b_\ell$. Otherwise $\ell$ could profitably deviate by claiming to be $h$ and quitting when beliefs drift down from initial beliefs by $b_\ell$. In this way, $\ell$ is able to maintain the same quitting threshold as truthfully declaring himself to be $\ell$ while also achieving a lower approval threshold $B_h$.

Theorem 2 assumes that $DIC(Z_h, Z_\ell)$ is slack and $DIC(Z_\ell, Z_h)$ is binding. This will not always be the case: there are parameter values for which $DIC(Z_h, Z_\ell)$ must bind. This comes about from the fact that $R$ may have to give $\ell$ a low initial threshold to induce any experimentation. Since $h$ has a higher belief than $\ell$ after observing $X_t$, $h$ (when reporting to be $\ell$) will still have positive continuation value when in $\ell$’s incentivization regime, creating incentives for $h$ to imitate $\ell$. However, we can show that if $Z_h$ is high enough, then the incentives of $R$ and $h$ are sufficiently aligned and $DIC(Z_\ell, Z_h)$ binding
A fast-track mechanism: the approval threshold for $h$ (the upper dashed line) starts off low, but, when $X_t$ crosses the failure threshold (the initial lower dotted line at $X = -0.22$), the approval threshold jumps up and drifts down as $M_t^X$ decreases.

is sufficient for $DIC(Z_h, Z_\ell)$ to be slack.

**Proposition 4.** If $Z_h$ is high enough, then the mechanism from Theorem 2 satisfies $DIC(Z_h, Z_\ell)$.

Turning to the situation in which $DIC(Z_h, Z_\ell)$ binds, it is not hard to adapt our previous proof to show that the $h$ type mechanism maintains the same main qualitative features as before, with a couple of additional wrinkles to provide additional incentives for $h$ to declare his type truthfully. Considering $\ell$'s mechanism, it is intuitive that as long as $\ell$ finds it optimal to experiment, $h$ will as well (since $h$ is more optimistic about the state being $H$). However, this could create a perverse incentive for $R$ to approve only after drops in the belief (since this rewards $\ell$ in expectation more than $h$). We will show that we can rule out such mechanisms when $Z_\ell$ is low enough\(^{13}\) and verify that the optimal mechanism will look very similar to that of Theorem 2.

**Theorem 3.** For each type $i$, define a mechanism $\tau_i = \inf\{t : X_t \not\in (b_i, B^i(M_t^X))\}$ and $d^i_\tau = 1(X_\tau = B^i(M_t^X))$ where

\(^{13}\)The assumption that $Z_\ell$ is low is reasonable in many applications. In the case of drug approval, over 90% of all drugs that begin a clinical trial fail to be approved (see FDA (2017)).
\[ B^\ell(M^X_t) = \begin{cases} B^\ell_1 & M^X_t \in [b^- \lor b^*_\ell(B^\ell_1), 0), \\ B^\ell_t(M^X_t) & M^X_t \in [b^-, b^- \lor b^*_\ell(B^\ell_1)), \end{cases} \]

\[ B^h(M^X_t) = \begin{cases} B^h_1 & \text{if } M^X_t \in [b^h_1, 0], \\ B^h_t(M^X_t) & \text{if } M^X_t \in [b^*_h(B^h_2), b^h_1], \\ B^h_2 & \text{if } M^X_t \in [b^*_h(B^h_2), b^*_\ell(B^h_2)), \\ B^h_h(M^X_t) & \text{if } M^X_t \leq b^*_h(B^h_2), \end{cases} \]

for some \((B^\ell_1, b^-)\) and \((B^h_1, B^h_2, b^h_1)\) such that \(b^- = b^- \lor b^*_h(B^\ell_1)\) and \(b^h_1 = -\frac{\sigma}{\phi} Z_h\). When \(DIC(Z_h, Z_\ell)\) is binding, such a mechanism is optimal for \(h\). It is also optimal for \(\ell\) if \(a < 0\) and \(Z_\ell < \log(-a)\).

There are two main differences between the mechanism when \(DIC(Z_h, Z_\ell)\) is binding and when it is slack. When it binds, \(R\) may reject \(\ell\) early (i.e. \(b^- > b^-(0)\)) in order to lower incentives for \(h\) to imitate \(\ell\). Additionally, it may be that the second stationary regime for \(h\) starts below \(R\)'s symmetric information solution (so that \(R\) can provide additional incentives for \(h\) while maintaining \(\ell\)'s incentive to quit).

Interestingly, unlike \(h\)'s mechanism when \(DIC(Z_\ell, Z_h)\) is binding, \(\ell\)'s mechanism does not qualitatively change much even when \(DIC(Z_h, Z_\ell)\) is binding. The fast-track feature of \(h\)'s mechanism comes from the backloading of punishments by \(R\), whereas the distortions in \(\ell\)'s mechanism from \(DIC(Z_h, Z_\ell)\) come through in early rejection of the project.

Our qualitative analysis of the optimal mechanism leaves us with very few parameters over which we must optimize. For high enough \(Z_h\), the optimal mechanism is completely pinned down by the choice of the thresholds of the stationary regime \((B^h_1, b^h_1)\) and \((B^\ell_1, b^-)\). When \(DIC(Z_h, Z_\ell)\) is binding, we have consider three parameters each for \(h, \ell\): \((B^h_1, b^h_1, b^h_2)\) and \((B^\ell_1, b^\ell_1, b^-)\). This reduction makes the problem computationally tractable since the choice of these thresholds will pin down the rest of the mechanism.

In Appendix H we show how to write down the problem in a way that can be solved quantitatively.

### 5.2 Comparative Statics

While all of the previous section allows for general \(a\), we now restrict attention to the simplest asymmetric information case when \(a = 1\) (so that \(A\) always prefers immediate approval). In this case, the misalignment of \(R\) and \(A\)'s preferences is particularly severe,
making it more difficult for \( R \) to elicit \( A \)’s private information. As we will see, this difference in preferences opens up a number of interesting comparative statics.

In the symmetric information case, increasing the cost \( c \) unambiguously hurts \( R \), since it makes it more difficult for \( R \) to incentivize experimentation. However, with asymmetric information this is no longer the case. Additional costs may be of use as a screening device. When \( c \) becomes small, it becomes increasingly harder for \( R \) to induce \( \ell \) to quit while still inducing \( h \) to keep experimenting. Taking the limit as \( c \to 0 \), we get that the private information of \( A \) is not used at all.

**Proposition 5.** As \( c \to 0 \) the optimal mechanisms for \( h, \ell \) converge to value of the single-decision maker problem for \( R \) with prior \( P(Z_h)\pi_h + (1 - P(Z_H))\pi_\ell \).

With asymmetric information and the absence of monetary transfers, costly experimentation provides a tool for screening of types, as detailed in the following proposition. This result can speak to the debate on who should fund drug trials (drug companies or government agencies), providing a reason for requiring the companies by requiring them to have some “skin in the game” and making it easier to elicit any private information the companies may have.

**Proposition 6.** The value of the optimal mechanism is non-monotonic in \( c \) under asymmetric information while the value of the optimal mechanism is strictly decreasing in \( c \) under symmetric information.

To illustrate the idea behind this, consider the case of \( \pi_h \approx 1 \) and \( \pi_\ell \approx 0 \). As \( c \to 0 \), screening becomes impossible and \( R \) essentially gets the value of the symmetric information mechanism. However, as \( c \) becomes large, it becomes possible to screen out \( \ell \) from \( h \) with a very low approval threshold, thereby increasing the speed of approval and increasing \( R \)’s utility.

Additionally, we might wonder whether or not it is beneficial to \( R \) for \( A \) to have private information about \( \theta \). On one hand, if \( R \) can make use of \( A \)’s information, then it is beneficial to \( R \). On the other hand, private information introduces information rents and can add distortions into \( R \)’s optimal mechanism. Which effect is greater is not ex-ante obvious. To answer this question, we compare the case of symmetric information to the case in which \( A \) has perfect information about \( \theta \). The following proposition shows that asymmetric information is in fact better for \( R \).

**Proposition 7.** Let \( \pi_0 \) be the prior of \( R \). Then the value to \( R \) of optimal mechanism under asymmetric information in which \( A \) learns \( \theta \) perfectly is higher than the value to \( R \) of the optimal mechanism under symmetric information with prior \( \pi_0 \).
6 General Markov Process

It is natural to wonder how the results of the model depend on the particular framework used. For example, how does the payoff structure determine the optimal mechanism? Does the exact specification of the diffusion process qualitatively determine the optimal strategy? To answer these questions, we generalize the symmetric information model to allow for a wide range of utility functions and stochastic processes and show that the optimal mechanism retains the same form as in Section 4.1.

Let \( X_t \) be a one-dimensional diffusion process on \( I \subseteq \mathbb{R} \) (where \( I \) is the interval \([a, a] \) and \( a, a \) are possibly \(-\infty, \infty \) respectively) which solves the stochastic differential equation

\[
dX_t = \mu(X_t)dt + \sigma(X_t)dB_t, \tag{1}
\]

for some Borel functions \( \mu : I \to \mathbb{R} \), \( \sigma : I \to \mathbb{R} \) (with \( \sigma(X_t) \) uniformly bounded above 0) and a given \( X_0 \). We assume \( \mu, \sigma \) are such that equation 1 has a unique (weak) solution.

For an arbitrary mechanism \((\tau, d_\tau)\), let \( R \)'s utility be given by \( \mathbb{E}[e^{-r\tau}g(X_\tau, d_\tau)|X_0] \) and \( A \)'s utility be given by \( \mathbb{E}[e^{-r\tau}f(X_\tau, d_\tau)|X_0] \). As before, we assume that \( A \) cannot sign binding long-term contracts (one-sided commitment). If \( A \) chooses to quit at \( X_t \), he receives his outside option, \( f(X_t, 0) \), which is equal to his payoff from rejection and may depend on \( X_t \). The mechanism design problem for \( R \) can then be written as

\[
\begin{align*}
[\text{GSM}] : & \quad \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau}g(X_\tau, d_\tau)|X_0] \\
\text{subject to} & \quad \mathbb{E}[e^{-r(\tau \wedge \tau')}f(X_{\tau \wedge \tau'}, d_\tau 1(\tau \leq \tau'))|X_0] \leq \mathbb{E}[e^{-r\tau}f(X_\tau, d_\tau)|X_0]
\end{align*}
\]

In order to use the arguments as sketched in Section 4.1, we need to place several assumptions (natural in many commonly studied environments) on the \( f, g \).

**Assumption 2.** We assume that for \( w \in \{f, g\} \), the following are satisfied:

- **1. Pure delay is sub-optimal:** For \( \alpha \in \{0, 1\} \) and \( b < B \), we have
  \[
  \mathbb{E}[e^{-r(\tau \geq (B) \wedge \tau(b))}w(X_{\tau \geq (B) \wedge \tau(b)}, \alpha)|X_t] < w(X_t, \alpha).
  \]

- **2. Decision Threshold:** If \( X' > X \), then \( w(X, 1) \geq w(X, 0) \) implies \( w(X', 1) > w(X', 0) \).
3. Continuity: $w(X, \alpha)$ is continuous in $X$ for each $\alpha \in \{0, 1\}$.

4. There exists $X$ such that for any $X_0 < X$, immediate rejection solves
\[
\sup_{(\tau, d_\tau) \in \Delta_C} \mathbb{E}[e^{-r\tau} g(X_\tau, d_\tau)|X_0] \text{ for } \Delta_C = \{(\tau, d_\tau) : \mathbb{E}[e^{-r\tau} f(X_\tau, d_\tau)|X_0] \geq f(X_0, 0)\}.
\]

5. $A$ prefers approval whenever $R$ does: $g(X, 1) > g(X, 0) \Rightarrow f(X, 1) > f(X, 0)$.

6. $\lim_{x \to \infty} \mathbb{E}[e^{-r\tau(x)}|y]w(x, 1) = 0$ for any $y \in I$.

While the list of assumptions may seem long, each part of the assumption is generally very mild and will fit many other models besides the one we have analyzed so far. Parts 1 – 3 are straightforward and are satisfied in most stopping problems considered in the literature. Part 2 implies that there are myopic cutoff points for each player, above which they prefer approval and below which they prefer rejection, which we denote $X^R_{my}, X^A_{my}$ respectively-i.e., $g(X^R_{my}, 1) = g(X^R_{my}, 0)$ (and define $X^A_{my}$ similarly for $f$). Part 4 ensures that there is a lower bound below it is too costly for $R$ to incentivize $A$ to continue. Part 5 ensures that $R$ has an incentive to stop at some point if $X_\tau$ goes too low and Part 6 is merely a technical condition needed to ensure payoffs do not diverge.

Next, we need to place some assumptions on the preferences of the players in relation to their optimal approval thresholds. As before, we will look at situations in which $A$ prefers “less” experimentation (i.e., a lower approval threshold) than $R$ does. Define $A$’s utility to static threshold mechanism which stops whenever $X_\tau \notin [b, B]$ (so $\tau = \tau_\geq(B) \land \tau_\leq(b)$) and takes approves only at $B$:
\[
\hat{V}(B, b, X) = \mathbb{E}[e^{-r\tau} f(X_\tau, 1(X_\tau = B))|X],
\]
We define a similar but extended version for $R$, $\hat{J}$, in which we allow for $R$’s continuation value of $u$ at $b$:
\[
\hat{J}(B, b, X, u) = \mathbb{E}[e^{-r\tau}(1(X_\tau = B)g(X_\tau, 1) + 1(X_\tau = b)u)|X].
\]

Let us suppose that the optimal mechanism were restricted to reject at $\tau(b)$. We can then ask what the preferred static approval threshold would be for $R$ and $A$. As in Section 4, we will want that $R$ prefers a higher approval threshold than $A$ does. We define $B^*_R(b), B^*_A(b)$ to be functions which give us the optimal approval thresholds of $R$ and $A$ respectively:
\[
B^*_R(b) := \arg\max_B \mathbb{E}[e^{-r(\tau_\geq(B) \land \tau(b))} g(X_\tau, 1(X_\tau = B))|X_0]
\]
\[
B^*_A(b) := \arg\max_B \mathbb{E}[e^{-r(\tau_\geq(B) \land \tau(b))} f(X_\tau, 1(X_\tau = B))|X_0]
\]
This allows us to formally state our assumption that $R$ prefers a higher threshold. We assume that the utility with respect to a static approval threshold is strictly single-peaked and that $R$’s preferred approval threshold is greater than that of $A$.

**Assumption 3.** For $u \geq g(b,0)$, both $\hat{V}(B,b,X)$ and $\hat{J}(B,b,X,u)$ are strictly single-peaked in $B$ and continuous in all arguments. When $u = g(b,0)$, $R$’s preferred approval threshold is higher than that of $A$: $B_R^*(b) > B_A^*(b)$.

We can then write down functions analogous to $b^*, B$ as defined previously:

\[
b^*(B) := \arg \max_b \hat{V}(B,b,X_0)
\]
\[
\tilde{B}(b,X) := \min \{B > \max_b B_A^*(b) : \hat{V}(B,b,X) = f(X,0)\}
\]
\[
B(X) := \lim_{\delta \to 0} \tilde{B}(X + \delta, X).
\]

We now make our final assumption, a condition on $\tilde{B}$, which will be used to ensure that $B$ exists.

**Assumption 4.** $\tilde{B}(b,X)$ is continuously differentiable.

With this in hand, we can show that state a generalization of Theorem 1 for general payoff functions and stochastic processes.

**Theorem 4.** Under Assumptions 2-4, the solution to GSM is given by $\tau = \inf \{t : X_t \notin (b,B(M_t^X))\}$ (for some $b$) and $d_\tau = 1(X_t \geq B(M_t^X))$ where the approval threshold is given by

\[
B(M_t^X) = \begin{cases} B^1 & \text{if } M_t^X \geq b^*(B^1) \\ B(M_t^X) & \text{if } M_t^X \leq b^*(B^1) \end{cases}
\]

Theorem 4 illustrates how we can expand the results of Section 4.1 to more general payoff structures (e.g., allowing $A$ to have state-dependent utility or $R$ to bear some cost of experimentation) and allows us to easily state the optimal mechanisms for a number of standard environments outside of the experimentation/learning framework considered up until now. For example, we can consider a real-option game similar to that of Grenadier et al. (2016) (assuming that (for this and the other examples) the parameters are such that the conditions of Assumptions 2-4)
**Example 1.** Suppose that \( A \) is tasked with running a project and \( R \) is an outside investor who can invest in the project. The value of the project, if invested in, is given by \( X_t \), which solves the stochastic differential equation

\[
  dX_t = \mu X_t \, dt + \sigma X_t \, dB_t,
\]

The payoffs to \( R \) is zero if he rejects the project at \( X_t - K \) (for some \( K \in \mathbb{R}_+ \)) if he invests in the project and 0 if he does not while the payoff to \( A \) is \( \beta X_t + L_1 \) for some \( \beta \in \mathbb{R}_+ \) and \( L_1 \in \mathbb{R} \) if \( R \) invests and \( L_2 \in \mathbb{R}_+ \) if \( R \) rejects or \( A \) quits.

Our general model can also be used to model a manager’s decision of whether or not to promote an agent, in which case it is realistic for the agent’s outside option may depend on the beliefs about his type (something not capture in the model of Section 3).

**Example 2.** Suppose that \( R \) is a manager deciding whether or not to promote an agent \( A \). The agent pays a flow cost \( c \) until he is either promoted or let go. The agent’s type is either \( \{\theta_h, \theta_l\} \) and both learning and \( R \)’s payoffs are the same as in our base model. Both \( R \) and \( A \) have the same belief about the agent’s type.\(^{14}\) If \( A \) is promoted, he receives a payoff of 1 while if he is let go or he quits, he receives a payoff of \( f(X_t, 0) \) (we can interpret this as the outside wage he will receive given the market’s belief about his type, which may depend on the information revealed over the course of the game). If \( R \) promotes a \( \theta \) type, she receives a utility \( g_{\theta}(X_t, 1) \) (we allow her utility to depend on the agent’s type as well as her beliefs at the time of promotion; this can capture situations in which how the \( A \) is viewed by other employees at the time of promotion determines the payoff to approving \( A \)).

We can then also apply the model to study a lobbying situation. The decreasing threshold then corresponds to a gradual decrease in \( R \)’s demands, something we naturally see in many real life negotiations.

**Example 3.** Let \( R \) be a company lobbying with a politician \( A \) over the supply of some good. The politician wants to be seen as proactive and derives a utility of 1 whenever a deal is made. The degree of public support \( X_t \) determines the payment \( A \) can offer \( R \) for the good (let’s assume when public support is \( X_t \), \( A \) can offer \( R X_t \)). \( R \) derives a utility \( u_R(X_t) - K \) when a deal is reached and \( A \) offers the maximum possible when public

\(^{14}\)While it is often the case in real-life that \( A \) will have more information about his type than \( R \), we can view the type here as being indicative of the productivity match between \( R \) and \( A \), in which case the symmetric information assumption is more innocuous.
support is $X_t$. Additionally, $A$ can exert costly effort (with a flow cost $c$) to rally public support for the deal, so that public support evolves according to the diffusion process

$$dX_t = \mu(X_t)dt + \sigma(X_t)dB_t.$$ 

We can then interpret “approval” as $R$ agreeing to a deal when offered $X_t$. The approval threshold is analogous to the demand of $R$ and we see that $R$’s optimal negotiating strategy is to slowly decrease his demand if public support decreases.

7 Conclusion

In this paper, we present a model of a hypothesis testing problem with agency concerns. The optimal mechanism features a history dependent approval threshold, yet can still be solved for in a tractable way and can be written as a function of the minimum of the belief process. We find that the optimal mechanism when the agent posses no private information takes the form of a monotonically decreasing approval threshold. We are able to fully characterize the solution in the problem and show how these results can be generalized to a large class of payoff functions and diffusion processes, allowing us to explore their implications in a number of other economic settings, such as promotion or lobbying models.

We also apply the model to the case when the agent has private information, which adds to qualitatively features to the optimal mechanism. The optimal solution may take the form of a fast-track mechanism: high types are offered a low starting approval threshold, but if the evidence becomes too unfavorable, the approval threshold jumps up, entering a punishment phase in which it drifts back down slowly. These results show how agency problems may lead to an evolving and history-dependent approval rule.

References


Appendices

A Properties of $\bar{V}, \bar{J}$

Lemma 2. $\bar{V}(B,b,X), \bar{J}(B,b,X)$ are single-peaked in $B$ and, for a fixed $b$, we have $\argmax_B \bar{J}(B,b,X) \geq \argmax_B \bar{V}(B,b,X)$.

Proof. The single-peaked property follows from Lemma 1 of Chan et. al (2016). We argue that for a fixed $b$, the optimal threshold $B$ for $A$ is lower than that of $R$. For $R$, the optimal threshold solves the first-order condition:

$$
\Psi_B(1-e^{-B}) + \Psi e^{-B} = 0,
$$

and the optimal threshold for $A$ satisfies

$$
\Psi_B(1 + \frac{c}{r} + (a + \frac{c}{r})e^{-B}) - (a + \frac{c}{r})\Psi e^{-B} + \psi_B \frac{c}{r} = 0.
$$

We note that the derivative of the first-order condition with respect to $a$ is

$$(\Psi_B - \Psi)e^{-B} < 0,$$

and thus by the implicit function theorem and the second-order condition, the optimal threshold is decreasing in $a$. Therefore it is enough to prove the claim for $a = -1$, in which case $A$’s first-order condition is

$$
\Psi_B(1 + \frac{c}{r} + (-1 + \frac{c}{r})e^{-B}) - (-1 + \frac{c}{r})\Psi e^{-B} + \psi_B \frac{c}{r} = 0. \tag{2}
$$

Let $\Delta = B - b$. The derivative of the first-order condition with respect to $c$ is

$$
\Psi_B(1 + e^{-B}) - \Psi e^{-B} + \psi_B
$$

$$
= \Psi(\frac{R_1 e^{-R_1 \Delta} - R_2 e^{-R_2 \Delta}}{e^{-R_1 \Delta} - e^{-R_2 \Delta}})(1 + e^{-B}) - e^{-B} + \frac{R_2 e^{-\Delta} - R_1 e^{-\Delta}}{e^{-R_1 \Delta} - e^{-R_2 \Delta}}
$$

Thus the above is negative if and only if

$$
(R_1 e^{-R_1 \Delta} - R_2 e^{-R_2 \Delta})(1 + e^{-B}) - e^{-B}(e^{-R_1 \Delta} - e^{-R_2 \Delta}) + R_2 e^{-\Delta} - R_1 e^{-\Delta} \tag{3}
$$
At $\Delta = 0$, this is equal to

$$e^{-B}(R_1 - R_2) < 0.$$  

If we take the derivative of equation 3 with respect to $\Delta$ when equation 3 is equal to zero, we have

$$(-R_1^2 e^{-R_1\Delta} + R_2^2 e^{-R_2\Delta})(1 + e^{-B}) + e^{-B}(R_1 e^{-R_1\Delta} - R_2 e^{-R_2\Delta}) - R_2 e^{-\Delta} + R_1 e^{-\Delta}$$

$$= (R_1(1 - R_1)e^{-R_1\Delta} + R_2(R_2 - 1)e^{-R_2\Delta})(1 + e^{-B}) + e^{-B}((R_1 - 1)e^{-R_1\Delta} - (R_2 - 1)e^{-R_2\Delta})$$

$$= R_1 R_2(e^{-R_1\Delta} - e^{-R_2\Delta})(1 + e^{-B}) + e^{-B}e^{-B}((R_1 - 1)e^{-R_1\Delta} - (R_2 - 1)e^{-R_2\Delta})$$

$$< 0.$$  

Therefore, we know that equation 3 is always negative. Therefore, increasing $c$ decreases the right-hand side of equation 2. By the implicit function theorem and the second-order condition, we have that $A$’s optimal $B$ is decreasing in $c$. Since the $R$ and $A$ optimal thresholds are equal when $c = 0$, it must be that $A$’s optimal $B$ is lower than that of $R$.  

Lemma 3. $\tilde{V}$ satisfies single crossing of 0 with respect to $X$.  

Proof. Suppose that $\exists X^1 < X^2$ such that $\tilde{V}(B, b, X^1) = \tilde{V}(B, b, X^2) = 0$. Then for any $X \in (X^1, X^2)$, we have

$$\tilde{V}(B, b, X) = E[e^{-r\tau}(d_r + \frac{c}{r})X] - \frac{c}{r}$$

$$= E[e^{-r\tau}\mathbb{1}(\tau \leq X^1 < \tau \geq X^2)](\tilde{V}(B, b, X^1) + \frac{c}{r})X$$

$$+ E[e^{-r\tau}\mathbb{1}(\tau \geq X^1 > \tau \geq X^2)](\tilde{V}(B, b, X^2) + \frac{c}{r})X - \frac{c}{r}$$

$$= E[e^{-r\tau}\mathbb{1}(\tau \leq X^1 < \tau \geq X^2)]\frac{c}{r}X + E[e^{-r\tau}\mathbb{1}(\tau \geq X^1 > \tau \geq X^2)]\frac{c}{r}X - \frac{c}{r}$$

$$< 0.$$  

B General Optimal Stopping Properties

We now present several general properties of single-decision optimal stopping problems which will prove useful in our analysis.
Lemma 4. Let $Z_t$ be a solution to $dZ_t = \mu(Z_t)dt + \sigma(Z_t)dW_t$, where $W_t$ is a standard Brownian motion. Then for the problem

$$\sup_{(\tau, d_\tau)} \mathbb{E}[e^{-rt}(d_\tau g_1(Z_\tau) + (1 - d_\tau)g_2(Z_\tau))|Z_0].$$

There exists a solution of the form $\tau = \inf\{t : Z_t \notin (Z_r, Z_a)\}$ for $Z_i = Z_a$ or $Z_i = Z_r$.

Proof. We can note that conditional on stopping, it will be optimal to choose $d_\tau = 1 \iff g_1(Z_\tau) \geq g_2(Z_\tau)$. We can define $g(Z_\tau) = \max\{g_1(Z_\tau), g_2(Z_\tau)\}$ and rewrite the optimal problem as

$$\sup_{(\tau, d_\tau)} \mathbb{E}[e^{-rt}g(Z_\tau)|Z_0].$$

Because the process $Z_t$ is Markov and we have exponential discounting (and hence time consistency), the principle of optimality tells us that $Z_t$ is a sufficient state variable for the optimal policy from time $t$ onward.

Let us define the value function when current beliefs are $Z$ as

$$U(Z) := \sup_{\tau} \mathbb{E}[e^{-rt}g(Z_\tau)|Z].$$

As is standard, we can describe $\tau$ be a continuation region $C = \{Z : U(Z) > g(Z)\}$ and a stopping region $D = \{Z : U(Z) = g(Z)\}$. Although the continuation region could take a non-interval form (e.g., $C = [Z_1, Z_2] \cup [Z_3, Z_4]$ where $Z_1 \leq Z_2 \leq Z_3 \leq Z_4$), we are only concerned with the continuation region around $Z_0$. Since the diffusion process is continuous, for any $C$ which depends only on $Z$, there is another continuation region $C' = (Z_1', Z_2')$ which delivers the same expected value when starting at $Z_0$ (where $Z_1' = \sup_{(\tau, d_\tau)} Z \in \partial C : Z \leq Z_0$) is the highest boundary of $C$ which is below $Z_0$ and $Z_2' = \inf_{Z \in \partial C : Z \geq Z_0}$ is the lowest boundary point of $C$ above $Z_0$). Therefore, there is an optimal stopping policy in the form of a threshold strategy around $Z_0$. 

Lemma 5 (Duality). Let $\{\phi_i\}_{i=1}^n$ and $\Phi$ be bounded $\mathcal{F}_t^X$-measurable functions and define

$$C := \{(\tau, d_\tau) : \mathbb{E}[\phi_i(\tau, \omega, d_\tau)|Z_0] \leq 0 \ \forall i = 1, ..., n\}.$$

Suppose that $\exists(\tau, d_\tau)$ such that $\mathbb{E}[\phi_i(\tau, \omega, d_\tau)|Z_0] < 0 \ \forall i = 0, ..., N$ and that the optimal solution to $\sup_{(\tau, d_\tau) \in C} \mathbb{E}[\Phi(\tau, \omega, d_\tau)|Z_0]$ is such that $P(\tau > 0) = 1$. Then there is no duality gap-i.e.,
\[
\sup_{(\tau,d_\tau) \in C} E[\Phi(\tau,\omega,d_\tau)|Z_0] = \inf_{\lambda \in \mathbb{R}^{N+1}} \sup_{(\tau,d_\tau) \in C} E[\Phi(\tau,\omega,d_\tau)|Z_0] + \sum_{i=0}^{N} \lambda_i E[\phi_i(\tau,\omega,d_\tau)|Z_0].
\]

Moreover, the infimum is obtained by some finite \( \lambda^* \in \mathbb{R}^{N+1} \). Additionally, \((\tau,d_\tau)\) is a solution to \( \sup_{(\tau,d_\tau) \in C} E[\Phi(\tau,\omega,d_\tau)|Z_0] \) if and only if it is a solution to \( \sup_{(\tau,d_\tau)} E[\Phi(\tau,\omega,d_\tau)|Z_0] + \sum_{i=0}^{N} \lambda_i^* E[\phi_i(\tau,\omega,d_\tau)|Z_0] \) and complementary slackness conditions hold:

\[
\forall i, \lambda_i \cdot E[\phi_i(\tau,\omega,d_\tau)|Z_0] = 0.
\]


Lemma 6. Let \( G(\pi_t,d_t) = d_t(\alpha_1 \pi_t + \alpha_2) + \alpha_3 \) and \( \alpha_3 \geq 0 \). If the solution to

\[
V(\pi_t) = \sup_{(\tau,d_\tau)} E[e^{-r\tau}G(\pi_t,d_\tau)|\pi_0],
\]

is a static threshold mechanism with approval at both the upper threshold \( \pi_B \) and the lower threshold \( \pi_b \), then \( \pi_B = \pi_b \).

Proof. By standard arguments, \( V \) solves the differential equation \( r V(\pi) = \phi \pi^2 (1 - \pi^2) V''(\pi) \). Then \( V(\pi) > 0 \), which implies that \( V''(\pi) \geq 0 \). Let \( \beta = \frac{\pi - \pi_b}{\pi_B - \pi_b} \). Then

\[
\begin{align*}
\alpha_1 \pi_0 + \alpha_2 + \alpha_3 &\leq V(\pi_0) = V(\beta \pi_b + (1 - \beta) \pi_B) \\
&\leq \beta V(\pi_b) + (1 - \beta) V(\pi_B) \\
&= \beta (\alpha_1 \pi_b + +\alpha_2 + \alpha_3) + (1 - \beta) (\alpha_1 \pi_B + +\alpha_2 + \alpha_3) \\
&= \alpha_1 \pi_0 + \alpha_2 + \alpha_3,
\end{align*}
\]

which implies that immediate approval is optimal. \(\square\)

C Symmetric Information with One-Sided Commitment

Lemma 7. Any mechanism \((\tau,d_\tau)\) which satisfies all dynamic participation constraints must satisfy \( DP \). For any mechanism \((\tilde{\tau},\tilde{d}_\tau)\) which satisfies \( DP \), there exists another mechanism which satisfies all dynamic participation constraints and yields the same payoff as \((\tilde{\tau},\tilde{d}_\tau)\).
C.1 Proof of Lemma 7

**Proof.** Let \((\tau, d_\tau)\) be a mechanism which satisfies all dynamic participation constraints. Suppose that it did not satisfy a DP constraint-i.e., \(\exists \tau'\) such that

\[
E[e^{-r(\tau \wedge \tau')}(d_\tau 1(\tau \leq \tau') + \frac{c}{r})|Z_0] - E[e^{-r(\tau \wedge \tau')}(d_\tau 1(\tau \leq \tau') + \frac{c}{r})|Z_0] < 0,
\]

\[
\Rightarrow E[e^{-r(\tau - \tau')}(d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) 1(\tau > \tau')|Z_0] < \frac{c}{r}.
\]

Therefore, there must be a history \(h_{\tau'}\) such that at \(\tau'\) we have a strictly negative continuation value, a contradiction of the fact that all dynamic participation constraints hold. Therefore all \((\tau, d_\tau)\) which satisfy dynamic participation constraints also satisfy DP constraints.

Next, let us consider a mechanism \((\tilde{\tau}, \tilde{d}_\tau)\) which satisfies DP. We will construct a new mechanism, \((\hat{\tau}, \hat{d}_\tau)\) which satisfies all dynamic participation constraints and gives the same payoff to \(R\) as \((\tilde{\tau}, \tilde{d}_\tau)\). If \((\tilde{\tau}, \tilde{d}_\tau)\) satisfies all dynamic participation constraints, we are done. If some dynamic participation constraints are violated, we claim that it must happen only on a zero probability set. For some small \(\epsilon\), define \(\Gamma\) to be the set of \(\omega\) with a history in the set \(\{h_t \in H_t : E[e^{-r(\tau \wedge \tau')}(d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_0] \leq \frac{c}{r} - \epsilon\}\) (which is the set of histories such that \(A\)'s continuation value is at least \(\epsilon\) worse than quitting immediately). Let us define \(\tau'\) to be the first time the history falls in this set. Then we know that

\[
E[e^{-r(\tau \wedge \tau')}(d_\tau 1(\tau \leq \tau') \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_0] = E[1(\Omega \setminus \Gamma)e^{-r(\tau \wedge \tau')}(d_\tau 1(\tau \leq \tau') \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_0]
\]

\[
+ E[1(\Gamma)e^{-r(\tau \wedge \tau')}(d_\tau 1(\tau \leq \tau') \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_0]
\]

\[
= E[1(\Omega \setminus \Gamma)e^{-r(\tau \wedge \tau')}(d_\tau 1(\tau \leq \tau') \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_0]
\]

\[
+ E[1(\Gamma)e^{-r(\tau \wedge \tau')}(d_\tau 1(\tau \leq \tau') \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_0].
\]

Because \((\tilde{\tau}, \tilde{d}_\tau)\) satisfies DP, we know that
which are not consistent with an \( \omega \).

By definition of \( \Gamma \), we have that for each \( h \),

\[
\text{Therefore, we can specify (} \hat{\omega} \text{) such that the constraints are violated has probability zero.}
\]

\[
\text{Therefore, it cannot be that } \Gamma \text{ has strictly positive probability and this must hold for all } \epsilon.
\]

\[
\text{Suppose that } (\hat{\tau}, \hat{d}_r) \text{ does violate some dynamic participation constraints. Then we know that the set } \omega \in \Gamma \text{ such that the constraints are violated has probability zero. Therefore, we can specify (} \hat{\tau}, \hat{d}_r) \text{ to be equal to } (\hat{\tau}, \hat{d}_r) \text{ on the set of all histories not which are not consistent with an } \omega \in \Gamma \text{ and } (\hat{\tau}, \hat{d}_r) \text{ to reject at the first time that a dynamic participation constraint is violated. With probability one, the outcome from (} \hat{\tau}, \hat{d}_r) \text{ is the same as } (\hat{\tau}, \hat{d}_r) \text{ and thus they yield the same payoffs.}
\]

\[
\text{C.2 Proof of Theorem 1}
\]

Theorem 1 is a special case of Theorem 4). We therefore defer the proof to Section F and only prove here the independence of the threshold (in belief space) from initial beliefs.

\[
\text{Lemma 8. The optimal approval threshold in belief-space } B_2(M_t^Z) \text{ is independent of } Z_0 \text{ and depends only on } M_t^Z := \min\{Z_s : s \leq t\}.
\]
Proof. Take two different initial beliefs $Z_1 > Z_2$ and mechanisms $(r^i, d^i_r)$ respectively. Suppose that the initial thresholds $B_1, B_2$ (under $Z_1, Z_2$ respectively) are not equal. Let $J_i(z)$ be the utility to $R$ under the mechanism when starting at $Z_i$ when the belief $z$ is reached for the first time. Because the mechanism for $Z_1$ satisfies DP at $\tau(2)$, it must be that $J_1(Z_2) \leq J_2(Z_2)$ (otherwise we could replace it with the continuation mechanism from $Z_1$). Consider a hybrid mechanism $(\tau_3, d^3_r)$ using approval threshold $B_1$ until $\tau(Z_2)$ and the continuation mechanism at $\tau(Z_2)$ from $(\tau_2, d^2_r)$ If $B_1 > B_2$, then replacing $J_1(Z_2)$ with $J_2(Z_2)$ would be admissible since it would relax DP constraints, which would imply that $J_1(Z_2) \geq J_2(Z_2)$; by our previous statement, this implies equality. Suppose that $B_2 > B_1$; we know that DP constraints are satisfied at $\tau_1(Z_1)$ under $(\tau_2, d^2_r)$ and lowering the threshold will add slack to the DP constraints, so this hybrid mechanism will satisfy all our constraints. Since $(\tau_3, d^3_r)$ cannot yield a strict improvement, we know that

\[
\mathbb{E}[e^{-r_{\tau_3}d^3_r e^{Z_r^*} - 1} (1 + e^{Z_r}) | Z_1] = \mathbb{E}[e^{-r_{\tau_1} \land \tau(Z_2)} (d^1_r e^{Z_r^*} - 1) \mathbbm{1}(\tau_1 \leq \tau(Z_2)) + \mathbbm{1}(\tau_1 \leq \tau(Z_2))J_2(Z_2) | Z_1] \\
\leq \mathbb{E}[e^{-r_{\tau_1} \land \tau(Z_2)} (d^1_r e^{Z_r^*} - 1) \mathbbm{1}(\tau_1 \leq \tau(Z_2)) + \mathbbm{1}(\tau_1 \leq \tau(Z_2))J_1(Z_2) | Z_1].
\]

Since $J_1, J_2$ are independent of the history to reach them, this implies that $J_1(Z_2) \geq J_2(Z_2)$. With our previous observation, we can conclude that $J_1(Z_2) = J_2(Z_2)$ and so it is without loss to replace $J_2(Z_2)$ with $J_1(Z_1)$ which has the same initial threshold in belief space.

\[\square\]

D Symmetric Information with No Commitment

D.1 Proof of Proposition 1

Proof. Suppose that $R$ uses the mechanism from the case of one-sided commitment $(\tau^*, d^*_r)$ and $A$ uses the following strategy: experiment until $\tau^*$, at which he immediately stop and never restart experimentation, and if experimentation has stopped before $\tau^*$, then he immediately restarts experimentation and keeps experimenting until $\tau^*$.

We claim that this is an equilibrium. To see this, let’s first consider the incentives of $R$ to deviate. Suppose that the equilibrium calls for $R$ to approve at time $\tau^*$. If she doesn’t approve, then the agent quits experimenting at time $\tau^*$ forever. Since no new learning occurs, $R$ has a strict incentive to approve immediately at $\tau^*$ since $Z_{\tau^*} > 0$. Suppose $R$ had a profitable deviation $\tau'$ such that $\tau' \leq \tau^*$ and there is some history such
that \( \tau' \) approves strictly sooner than \( \tau^* \). Then \( \tau' \) will not violate any DP constraints (approving sooner would only slacken the DP constraints), contradicting the optimality of \( \tau^* \). Therefore no such deviation can exist.

Next, we consider the incentives of \( A \) to deviate from the proposed equilibrium. Note that under the proposed approval rule, since all the DP constraints hold, \( A \) has no incentive to quit early. If he were to quit early, \( R \) would believe that \( A \) will restart experimenting immediately and therefore not find it optimal to approve. Moreover, \( A \) has an incentive to stop experimenting at \( \tau^* \) since he believes that \( R \) will approve immediately. In the off-path event that \( R \) doesn’t approve, \( A \) believes that \( R \) will approve in the next instant and has no incentive to restart experimentation since it is costly and will not increase the probability of approval. Since neither \( A \) nor \( R \) have an incentive to deviate, \((\tau^*, d^*_\tau)\) is indeed an equilibrium.

\[ \square \]

### E Asymmetric Information with One-Sided Commitment

Rather than directly prove Proposition 2, we instead solve a generalization of \( AM^{h,\gamma} \) which will be useful in proving Lemma 14. Fix an arbitrary \( \gamma \in \mathbb{R} \) (the proof of Proposition 2 follow from letting \( \gamma = 0 \)); we define the problem \( AM^{h,\gamma} \) as

\[
[AM^{h,\gamma}] : \sup_{(\tau,d_{\tau})} \mathbb{E}[e^{-r\tau} (d_{\tau} e^{Z_{\tau}(1 + \gamma)} + (a\gamma - 1) + \gamma c) | Z_h] \\
\text{subject to } DP(Z_h), \ DIC(Z_{\ell}, Z_h, V_{\ell}).
\]

Let \((\tau_{SM}, d_{\tau}^{SM})\) be the solution to \( SM^\gamma(Z_h) \) where \( SM^\gamma(Z_h) \) is the symmetric information problem with prior \( Z_h \) when \( R \)'s payoffs depend on \( \gamma \) as above. We define \( b_{SM,\gamma} \) to be the \( X_t \) such that, when \( \ell \) has belief \( Z_t = Z_{\ell} + \phi \sigma X_t \), he would quit immediately if \( R \) proposed \( h \)'s symmetric mechanism \( SM^\gamma(X_t) \) (the symmetric mechanism starting at \( X_t, M_t^X = X_t \)):

\[
b_{SM,\gamma} := \max \{X_t : \sup_{\tau'} \mathbb{E}[e^{-r(\tau_{SM} \wedge \tau')} (d_{\tau}^{SM} 1(\tau \leq \tau') e^{Z_{\tau}} + a + \gamma c) | X_t, Z_t = Z_{\ell} + \phi \sigma X_t] = c \}.
\]

We begin by proving solving a relaxed problem \( RAM^h_N(\gamma) \) defined as

\[ ^{15} R \] will never find it profitable to reject earlier
\[ [\text{RAM}^h_\gamma]^N_N] : \sup_{(\tau,d_\tau)} \mathbb{E}[e^{-r(\tau \wedge \tau(b^{SM}\gamma))}(d_\tau(b^{SM}\gamma)(\frac{e^{Z_\tau}(1 + \gamma) + (a\gamma - 1)}{1 + e^{Z_\tau}} + \gamma \frac{c}{r})\mathbb{1}(\tau \leq \tau(b^{SM}\gamma))}
+ \mathbb{1}(\tau > \tau(b^{SM}\gamma))SM^{h\gamma}(b^{SM}\gamma)|Z_h]
\]

subject to \( \forall X_i \in T_N \cup \{b^{SM}\gamma}\)

\[ \text{RDIC}_\ell(X_i) : \mathbb{E}[e^{-r(\tau \wedge \tau(X_i) \wedge \tau(b^{SM}\gamma))}(d(X_i)\frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})\mathbb{1}|Z_t] \leq V_\ell + \frac{c}{r}. \]

Define an analogous version of \( H_N(X_t) \) by

\[ [H^h_\gamma(X_t)] : \sup_{(\tau,d_\tau)} \mathbb{E}[e^{-r(\tau \wedge \tau(X_t) \wedge \tau(b^{SM}\gamma))}(d_\tau(b^{SM}\gamma)(\frac{e^{Z_\tau}(1 + \gamma) + (a\gamma - 1)}{1 + e^{Z_\tau}} + \gamma \frac{c}{r})\mathbb{1}(\tau \leq \tau(b^{SM}\gamma))}
+ \mathbb{1}(\tau > \tau(b^{SM}\gamma))SM^{h\gamma}(b^{SM}\gamma)|Z_t]
\]

subject to \( \forall X_i \in \{X_j \in T_N \cup \{b^{SM}\gamma\} : X_j < X_t\} \)

\[ \text{RPK}(0) : \mathbb{E}[e^{-r(\tau \wedge \tau(X_t) \wedge \tau(b^{SM}\gamma))}(d_\tau(X_t)\mathbb{1}(\tau \leq \tau(b^{SM}\gamma))\frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})\mathbb{1}|Z_t - \Delta Z] \leq \frac{c}{r}. \]

**Lemma 9.** \( \text{RAM}^h_\gamma^N \) is an upper-bound on \( AM^{h\gamma} \).

**Proof.** First, we claim that the continuation value for \( AM^{h\gamma} \) at \( \tau(b^{SM}\gamma) \) is bounded above by \( SM^{h\gamma} \). This is clear, since any mechanism admissible with respect to \( AM^{h\gamma} \) must satisfy all DP constraints when starting at \( b^{SM}(\gamma) \) and thus is admissible with respect to \( SM^{h\gamma} \). Moreover, since we have dropped some constraints, for any \((\tau,d_\tau)\) admissible with respect to \( AM^{h\gamma} \), we have that \((\tau \wedge \tau(b^{SM}\gamma),d_\tau(b^{SM}\gamma)) \) is admissible with respect to \( \text{RAM}^h_\gamma^N \). The result then follows immediately. \( \square \)

**Lemma 10.** The solution to \( \text{RAM}^h_\gamma^N \) is given by a stationary approval threshold until the first binding constraint \( X^1 \). At \( X^1 \), the continuation mechanism solves \( H^h_\gamma^N(X^1) \).

**Proof.** We face the new complication in solving \( \text{RAM}^h_\gamma^N \) in that the expectations are taken with respect to different priors, which makes the arguments from Lemma 16 inapplicable. However, we can note that
\[
E[e^{-r(\tau \land \tau(X_i))}(d_\tau(X_i)e^{Z_\tau} + \frac{a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_h - \Delta Z] = \frac{e^{Z_h - \Delta Z}}{1 + e^{Z_h - \Delta Z}}E[e^{-r(\tau \land \tau(Z_i))}(\mathbb{1}(\tau' \leq \tau(X_i))d'_\tau + \frac{c}{r})|\theta = H] + \frac{1}{1 + e^{Z_h - \Delta Z}}E[e^{-r(\tau \land \tau(Z_i))}(d_\tau(X_i) + \frac{c}{r})|\theta = H] \\
+ \frac{1}{1 + e^{Z_h - \Delta Z}}E[e^{-r(\tau \land \tau(Z_i))}(d_\tau(X_i) + \frac{c}{r})|\theta = L] = \frac{1 + e^{Z_h}}{1 + e^{Z_h - \Delta Z}}E[e^{-r(\tau \land \tau(Z_i))}(d_\tau(X_i) + \frac{c}{r})|\theta = L] \\
+ \frac{1}{1 + e^{Z_h - \Delta Z}}E[e^{-r(\tau \land \tau(Z_i))}(d_\tau(X_i) + \frac{c}{r})e^{-\Delta Z}e^{Z_\tau(X_i)} + a][Z_h].
\]

With this change of expectation, we can now apply Lemma 16 in order to conclude that the optimal mechanism consists of a static approval threshold until the first binding constraint \(X^1\) is reached or \(b^{SM,\gamma}\).

All that is left is to show that the continuation mechanism at \(\tau(X^1)\) is the solution to \(H_{X^1}^{b^{SM,\gamma}}(X^1)\). This is immediate if \(X^1 = b^{SM,\gamma}\). Therefore suppose that \(X^1 > b^{SM,\gamma}\). As noted in Lemma 16, the optimal mechanism from \(\tau(X^1)\) onward is independent of the history up to \(\tau(X^1)\). We know that for \(X_i < X^1\), we have

\[
E[e^{-r(\tau \land \tau(X_i))}(d_\tau(X_i)e^{Z_\tau} + \frac{a}{1 + e^{Z_\tau}} + \frac{c}{r})|X_0, Z_\ell] \leq V_\ell + \frac{c}{r}, \quad (4)
\]

Let \(\tau[h_\tau(X^1)](X_i)\) be the threshold quitting time of \(X_i\) given \(X_0 = X^1\). Since the mechanism at \(\tau(X^1)\) is independent of the history until \(\tau(X^1)\), we know that

\[
E[e^{-r(\tau \land \tau(X_i))}(d_\tau(X_i)e^{Z_\tau} + \frac{a}{1 + e^{Z_\tau}} + \frac{c}{r})|X_0, Z_\ell] = E[\mathbb{1}(\tau \leq \tau(X^1))e^{-r(\tau(X^1))}(d_\tau(X_i)e^{Z_\tau} + \frac{a}{1 + e^{Z_\tau}} + \frac{c}{r})|X_0, Z_\ell] \\
+ E[\mathbb{1}(\tau > \tau(X^1))e^{-r(\tau(X^1))}e^{-r(h_\tau(X^1)\land \tau(X^1))}(d_\tau(X_i)e^{Z_\tau} + \frac{a}{1 + e^{Z_\tau}} + \frac{c}{r})|X^1]|X_0, Z_\ell] = V_\ell + \frac{c}{r} - \frac{c}{r}E[e^{-r(\tau(X^1))}\mathbb{1}(\tau > \tau(X^1))|X_0, Z_\ell] \\
+ E[\mathbb{1}(\tau > \tau(X^1))e^{-\tau(X^1)}|X_0, Z_\ell]E[e^{-r(h_\tau(X^1)\land \tau(X^1))}(d_\tau(X_i)e^{Z_\tau} + \frac{a}{1 + e^{Z_\tau}} + \frac{c}{r})|X^1, Z_\ell],
\]

54
which, with equation 4, implies that

\[
\mathbb{E}[e^{-r(\tau_{[h,\gamma]}(X_1))}(d_r(X_i)\frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})] \leq \frac{c}{r},
\]

and thus the expected continuation value to \( \ell \) of continuing until \( \tau(X_i) \) is weakly negative at \( X^1 \). Therefore, the continuation mechanism at \( \tau(X^1) \) is admissible with respect to \( H_N^{h,\gamma}(X^1) \) and thus \( H_N^{h,\gamma}(X^1) \) is weakly higher than the value of the optimal mechanism at \( \tau(X^1) \). Suppose that using the mechanism \((\tau^H, d^H_r)\) which solves \( H_N^{h,\gamma}(X^1) \) yielded a strictly higher value to \( R \) than the continuation mechanism at \( \tau(X^1) \). Then replacing the continuation mechanism at \( \tau(X^1) \) with \((\tau^H, d^H_r)\) would give a value of

\[
\mathbb{E}[\mathbb{I}(\tau \leq \tau(X^1))e^{-r(\tau \wedge \tau(X^1))}(d_r(X_i)\frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})]|Z_h] + \mathbb{E}[\mathbb{I}(\tau > \tau(X^1))e^{-r(\tau(X^1))}H^h(\gamma, X^1)|Z_h]
\]

where \( Z_h = Z_\ell + \frac{\phi}{\sigma}X^1 \) is belief of \( \ell \) at \( X^1 \). Thus, switching to \((\tau^H, d^H_r)\) at \( \tau(X^1) \) is admissible with respect to \( RSM_N^{h,\gamma} \). Therefore, the optimal mechanism at \( \tau(X^1) \) must solve \( H_N^{h,\gamma}(X^1) \).
Lemma 11. The limit as $N \to \infty$ of the mechanisms which solve $H_N^{h,\gamma}(X^1)$ is given by 
\[
\tau = \inf \{ t : X_t \geq B_t(M_t^X) \} \quad \text{and} \quad d_\tau = \mathbb{1}(X_\tau = B_t(M_t^X)).
\]

Proof. Repeated application of Lemma 16 yields the conclusion that the approval rule is a sequence of thresholds $\{B^i_h\}$ such that the approval changes from $B^i_h$ to $B^{i+1}_h$ when the $i$th binding $RDIC$ constraint is reached. Applying Lemma 10, we know that at each binding constraint $X^i_j > b^{SM,\gamma}_h$, the continuation mechanism solves $H_N^{h,\gamma}(X^i_j)$. Let us consider the limit as $N \to \infty$. By the same arguments as in Lemma 18, we get that as long as $RPK$ constraints are binding, the upper approval threshold must converge to $B^\ell(M_t^X)$ as $N \to \infty$. All that is left to show is that $RPK$ constraints are binding until $M_t^X = b^{SM,\gamma}_h$. Consider a fixed $N$ and suppose that there exists $X^i_j > b^{SM,\gamma}_h$ such that $X^i_j$ is not the last constraint before $b^{SM,\gamma}_h$ (i.e., $\exists X^i_i \in T_N \in (b^{SM,\gamma}_h, X^i_j)$) and all $RPK$ constraints prior to $b^{SM,\gamma}_h$ are slack in $H_N^{h,\gamma}(X^i_j)$. In this case, the complementary slackness conditions in our Lagrangian amount to solving 

\[
\sup_{(\tau,d_\tau)} \mathbb{E}[e^{-r(\tau \wedge \tau(b^{SM,\gamma}))}(d_\tau(b^{SM,\gamma}) - (z_\tau(1 + \gamma) + (a\gamma - 1)) \frac{1}{1 + e^{z_\tau \gamma}} + \gamma \frac{c}{r}) \mathbb{1}(\tau \leq \tau(b^{SM,\gamma})) \\
+ \mathbb{1}(\tau > \tau(b^{SM,\gamma}))SM^{h,\gamma}(b^{SM,\gamma})(Z_t)].
\]

The solution to this consists of only an approval threshold at $B$ (since $SM^{h,\gamma}(b^{SM,\gamma}) \geq 0$, $R$ never benefits from rejecting prior to $b^{SM,\gamma}_h$). We claim that $B$ must be equal to the initial approval threshold in the symmetric mechanism $B^{1}_{SM,\gamma}$. If there was a $B < B_h(b^{SM,\gamma})$ which was not equal to $B^{1}_{SM,\gamma}$ and yielded a higher utility, then this $B$ will yield a higher utility in the symmetric information problem, since it only relaxes the $RDP$ constraints for $h$. Suppose then that there was a $B > B_h(b^{SM,\gamma})$ which yielded strictly higher utility. It is straightforward that this $B$ would be optimal in $h$’s symmetric mechanism, which implies that at $X^i$, the symmetric mechanism for $h$ must have a binding $DP$ constraint. However, this cannot be so since $b^{SM,\gamma}_h$ is above where $h$’s $DP$ constraints begin to bind. 

Lemma 12. Let $(\tau^N, d^N_\tau)$ solve $RAM_N^{h,\gamma}$ and $(\tau, d_\tau) := \lim_{N \to \infty} (\tau^N, d^N_\tau)$. Then $(\tau, d_\tau)$ solves $AM^{h,\gamma}$.

Proof. Since $RAM_N^{h,\gamma}$ is an upper-bound on $AM^{h,\gamma}$, all that is necessary is to verify that $(\tau, d_\tau)$ is admissible with respect to $AM^{h,\gamma}$. First, we need to show that $h$ has no incentive to quit early. This is immediate after $\tau(b^{SM,\gamma})$ since the symmetric mechanism, which satisfies all $DP$ constraints, is used. Before $\tau(b^{SM,\gamma})$, the continuation value to $h$ is always positive since $B_t(M_t^X) < B_h(M_t^X)$.
We also need to check that \( \ell \)'s DIC constraint truly does hold. To this end, we claim that the optimal quitting rule \( \ell \) could use is a threshold quitting rule. Note that the optimal mechanism is Markov with respect to \((X_t, M^X_t)\) and thus, by the principle of optimality, \( \ell \)'s optimal quitting rule will also be Markov with respect to \((X_t, M^X_t)\) and thus can be expressed as 
\[
\tau' = \inf\{t : X_t \geq \kappa(M^X_t)\} \land \tau(b_A) \text{ for some function } \kappa \text{ and constant } b_A \text{ (where we take } \kappa(M^X_t) > B(M^X_t) \text{ to imply never quitting at any } X_t \in (M^X_t, B(M^X_t))) . \]
However, since \( B(M^X_t) \leq B_{\ell}(M^X_t) \), it is never optimal \( \ell \) to quit at \( \kappa(M^X_t) > B(M^X_t) \). Therefore, the optimal quitting rule must be equivalent to 
\[
\tau' = \tau(b_A) \text{ for some } b_A \text{ (which is a threshold quitting contraint) and is included in our relaxed problem.}
\]
\[\square\]

E.1 Proof of Proposition 3

Proof. As before, we define a relaxed version of \( AM^\ell \) to be

\[
[RAM^\ell_N] : \sup_{(\tau,d\tau)} \mathbb{E}[e^{-r\tau}d\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z\ell]
\]
subject to \( \forall X_i \in T_N \)

\[
RDP_\ell(X_i) : \mathbb{E}[e^{-r(\tau \land \tau(X_i))}(d\tau \mathbb{1}(\tau < \tau(X_i)) \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z\ell] \leq \mathbb{E}[e^{-r\tau}(d\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z\ell]
\]

\[
PK_\ell(V\ell) : \mathbb{E}[e^{-r\tau}(d\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z\ell] \geq V\ell + \frac{c}{r}.
\]

We claim that the value of \( RAM^\ell_N \) is (for some value \( \gamma \)) equivalent to

\[
[G^\ell,\gamma_N] : \sup_{(\tau,d\tau)} \mathbb{E}[e^{-r\tau}(d\tau \frac{e^{Z_\tau}(1 - \gamma) + (-a\gamma - 1)}{1 + e^{Z_\tau}} - \gamma \frac{c}{r})|Z\ell]
\]
subject to \( RDP_\ell(X_i) \forall X_i \in T_N, X_i < X^1. \)

By Lemma 5, we know that the value of \( RAM^\ell_N \) is equal to
\begin{align*}
&\inf_{(\gamma,\lambda)\in \mathbb{R}^{N+2}} \sup_{(r,d_e)} \mathbb{E}[e^{-r\tau}(d_r e^{Z_r}(1-\gamma) + (-a\gamma - 1) - \gamma \frac{c}{r})|Z_\ell] + \gamma V_\ell \\
&+ \sum_{i=0}^{N} \lambda_e(X_i)(\mathbb{E}[e^{-r(\tau \wedge \tau(X_i))}(d_r 1(\tau < \tau(X_i)) e^{Z_r} + a) + \frac{c}{r})|Z_\ell] - \mathbb{E}[e^{-r\tau}(d_r e^{Z_r} + a) + \frac{c}{r})|Z_\ell]) \\
&= \inf_{\gamma \in \mathbb{R}} \inf_{\lambda_e \in \mathbb{R}^{N+1}} \sup_{(r,d_e)} \mathbb{E}[e^{-r\tau}(d_r e^{Z_r}(1-\gamma) + (-a\gamma - 1) - \gamma \frac{c}{r})|Z_\ell] + \gamma V_\ell \\
&+ \sum_{i=0}^{N} \lambda_e(X_i)(\mathbb{E}[e^{-r(\tau \wedge \tau(X_i))}(d_r 1(\tau < \tau(X_i)) e^{Z_r} + a) + \frac{c}{r})|Z_\ell] - \mathbb{E}[e^{-r\tau}(d_r e^{Z_r} + a) + \frac{c}{r})|Z_\ell]) \\
&= \inf_{\gamma \in \mathbb{R}} G^{\ell,-\gamma}_N + \gamma V_\ell.
\end{align*}

We can apply Theorem 4 to $G^{\ell,-\gamma}_N$, yielding the desired form of the optimal mechanism.

\hfill \Box

\subsection*{E.2 Proof of Theorem 2}

\textbf{Proof.} Suppose that $B^\ell_h = b^\ell_b(B^1_\ell)$. Because $\ell$ receives zero expected utility conditional on reaching $b^\ell_h$, $\ell$'s expected utility is given by $V(B^1_\ell, b^\ell_h, Z_\ell)$. First we want to show that $B^1_h \leq B^\ell_h$. Let's consider the case when $DIC(Z_\ell, Z_h)$ is binding. For the sake of contradiction, suppose that $B^1_h > B^\ell_h$. The utility that $\ell$ gets from claiming to be $h$ is bounded above by $\max_b \tilde{V}(B^1_h, b, Z_\ell)$. Because $b^\ell_h = b^\ell_b(B^1_\ell)$, then the utility $\ell$ gets from truthfully reporting his type is given by $\max_b \tilde{V}(B^1_b, b, Z_\ell)$. By Assumption 1 and the fact that $\tilde{V}$ is single-peaked in $B$, we know that for $b > b^{SM}$ we have $\max_b \tilde{V}(B^1_h, b, Z_\ell) < \max_b \tilde{V}(B^1_b, b, Z_\ell)$, which contradicts $DIC(Z_\ell, Z_h)$ binding. The rest of the theorem follows from the arguments in the text.

\hfill \Box

\subsection*{E.3 Proof of Proposition 4}

\textbf{Proof.} We note that as $Z_h \to \infty$, the probability of approval condition on the state begin $H$ approaches 1 in the optimal mechanism since $b^\ell_h(-\frac{c}{2}Z_h) \to -\infty$. The proof is then obvious once we note that when $\pi_h = 1$, the incentives of both $R$ and $A$ are perfectly aligned in that they are given by $\mathbb{E}[e^{-r\tau}|H]$ and $\mathbb{E}[e^{-r\tau}|H](1 + \frac{c}{2})$ respectively.

\hfill \Box
E.4 Proof of Theorem 3

We split this proof into two Lemmas:

**Lemma 13.** The optimal mechanism for \( \ell \) when \( \text{DIC}(Z_h, Z_\ell) \) is binding and \( Z_\ell < \log(-a) \) (assuming \( a < 0 \)) is given by a dynamic approval threshold \( B_\ell(M_t) \), which is defined as

\[
B^\ell(M^X_t) = \begin{cases} 
 B^\ell_1 & M^X_t \in [b_r \lor b^*_\ell(B^\ell_1), 0), \\
 B^\ell_2(M^X_t) & M^X_t \in [b_r, b_r \lor b^*_\ell(B^\ell_1)).
\end{cases}
\]

for some \((B^\ell_1, b_r) \in \mathbb{R}_2\).

**Proof.** Since \( h \) will always have a higher belief than \( \ell \), we can conjecture that \( h \) will never quit as long \( \ell \) still finds it optimal to experiment. This leads us to define a relaxed problem in which we assume that \( h \) will not quit early:

\[
[\text{RAM}^B_{\text{BIND}}] : \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau}d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_\ell]
\]

subject to \( \text{PK}_\ell(V_\ell), \text{RDP}_\ell(X_i) \ \forall X_i \in T_N \),

\[
\text{RDIC}_h(V_h) : \mathbb{E}[e^{-r\tau}(d_\tau \max\{\frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}}, 0\} + \frac{c_r}{r}) | Z_h] \leq V_h + \frac{c_r}{r}.
\]

where we place \( \max\{\frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}}, 0\} \) in \( h \)'s utility to incorporate the fact that \( R \) cannot force \( h \) to accept approval when \( h \) would prefer that \( R \) reject. Using the arguments of Proposition 3 and a similar change of expectation as in Lemma 10 we can see that the solution \((\tau^*, d^*_\tau)\) to \( \text{RAM}^B_{\text{BIND}} \) is a sequence of static stopping thresholds \( \{B^i\}_{i=1}^M \) (where \( M \) is the number of binding \( \text{RDP} \) constraints; if \( M = 0 \), there is a single stopping threshold \( B^1 \)) and which moves from \( B^i \) to \( B^{i+1} \) when the \( i \)th binding \( \text{RDP} \) constraint is reached, as well as a lower stopping threshold \( b_r \).

Because \( Z_\ell \) is assumed to be sufficiently low, approval at \( b_r \) only decreases both \( \ell \) and \( R \)'s utility. Since we assume that \( h \) always has the outside option of quitting when \( R \) approves, rejection at \( b_r \) would be strictly better than approval. Moreover, \( R \) must always approve at the upper-thresholds so as to maintain \( \text{RDP}_\ell(X_i) \). Taking the limit as \( N \to \infty \) and applying the same arguments as in our other cases gives us the optimal mechanism.

\[\square\]
Lemma 14. When $DIC(Z_h, Z_\ell)$ is binding, there exists $(b^1_h, B^1_h, B^2_h)$ such that the optimal mechanism for $h$ is given by $\tau = \inf \{ t : X_t \geq B^h(M^X_t) \} \wedge \tau(-\sigma_Z h)$

$$B^h(M^X_t) = \begin{cases} B^1_h & \text{if } M^X_t \in [b^1_h, 0], \\ B^2_h & \text{if } M^X_t \in [b^2_h(B^2_h), b^1_h], \\ B^1_h & \text{if } M^X_t \in [b^2_h(B^2_h), b^1_h(B^2_h)], \\ B^2_h & \text{if } M^X_t \leq b^2_h(B^2_h), \end{cases}$$

and $d_\tau = 1(X_\tau = B^h(M^X_t))$.

Suppose that $DIC(Z_h, Z_\ell)$ is binding and let $V_h$ be the utility promised to $h$ under the optimal mechanism. The design problem for $h$’s optimal mechanism is similar to that in Lemma 9 except for the addition of a promise keeping constraint for $h$:

$$[AM^h] : \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r \tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_h]$$

subject to $DP(Z_h), DIC(Z_\ell, Z_h, V_\ell)$

$$PK(V_h) : \mathbb{E}[e^{-r \tau} (d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_h] \geq V_h + \frac{c}{r}.$$  

Proof. We define a relaxed problem, dropping quitting rules except for threshold quitting rules in $T_N$ as

$$[RAM^h_N] : \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r \tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_h]$$

subject to $\forall X_i \in T_N$

$$RDP_h(X_i) : \mathbb{E}[e^{-r (\tau \wedge \tau(X_i))} (d_\tau(X_i) \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_h] \leq \mathbb{E}[e^{-r \tau} (d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_h],$$

$$RDIC_\ell(X_i) : \mathbb{E}[e^{-r (\tau \wedge \tau(X_i))} (d_\tau(X_i) \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_\ell] \leq V_\ell + \frac{c}{r},$$

$$PK(V_h) : \mathbb{E}[e^{-r \tau} (d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_h] \geq V_h + \frac{c}{r}.$$  

By the same arguments as in Proposition 3, $\exists \gamma < 0$ such that the solution to $RAM^h_N$ is the solution to
\[
\sup_{(r,d,\tau)} \mathbb{E}[e^{-r\tau}(d_r e^{Z_\tau} - 1) - \gamma(d_r e^{Z_\tau} + \frac{c}{r})|Z_h] + \gamma V_h
\]

subject to \(RDP_h(X_i)\forall X_i \in T_N\),

\[
RDIC(Z_\ell, Z_h) : \mathbb{E}[e^{-r(\tau \wedge \tau(X_i))}(d_r(X_i) e^{Z_\tau} + \frac{c}{r})|Z_\ell] \leq V_\ell + \frac{c}{r}.
\]

By applying Lemmas 9 - 11, we get our desired result for the optimal stopping mechanism (where \(B^2_h\) is the initial level set for \(SM^h(\gamma)\)) as \(N \to \infty\).

\[\square\]

E.5 Proof of Proposition 5

Proof. Consider the case of \(c = 0\). Let \(\alpha_i = \mathbb{E}[e^{-r\tau} \mathbb{1}(d_r^i = 1)|\theta = H]\) and \(\beta := \mathbb{E}[e^{-r\tau} \mathbb{1}(d_r^i = 1)|\theta = L]\) be the discounted probability of approval for type \(Z_i\) when \(\theta = H\) and \(\theta = L\) (respectively). In order to preserve incentive compatibility, we must have

\[
\pi_h \alpha_h + (1 - \pi_h) \beta_h \geq \pi_h \alpha_\ell + (1 - \pi_h) \beta_\ell
\]
\[
\pi_\ell \alpha_\ell + (1 - \pi_\ell) \beta_\ell \geq \pi_\ell \alpha_h + (1 - \pi_\ell) \beta_h.
\]

By optimality of \(\tau_h, \tau_\ell\), we also must have

\[
\pi_h \alpha_h - (1 - \pi_h) \beta_h \geq \pi_h \alpha_\ell - (1 - \pi_h) \beta_\ell
\]
\[
\pi_\ell \alpha_\ell - (1 - \pi_\ell) \beta_\ell \geq \pi_\ell \alpha_h - (1 - \pi_\ell) \beta_h.
\]

Adding the equations using \(\pi_\ell\), we get \(\alpha_\ell \geq \alpha_h\). Doing the same with \(\pi_h\), we get that \(\alpha_h \geq \alpha_\ell\). Therefore we must have \(\alpha_h = \alpha_\ell\) and therefore \(\beta_h = \beta_\ell\). Therefore, it is without loss to offer both types the same mechanism. This mechanism must maximize \(R\)'s utility subject to offering both types the same mechanism, which corresponds to \(R\)'s optimal solution with prior \(P(Z_h)\pi_h + P(Z_\ell)\pi_\ell\). A straightforward application of the Theorem of the Maximum yields the statement of the proposition.

\[\square\]
E.6 Proof of Proposition 6

Proof. Suppose that \( \pi_h = 1 \) and \( \pi_\ell = 0 \). We first examine a limiting case where the signal to noise ratio \( \frac{\mu}{\sigma} \to 0 \) and \( c \to 0 \). We claim that the value of the optimal mechanism is zero. By Proposition 5, we know that the value of the optimal mechanism converges to that of a single decision maker. As \( \frac{\mu}{\sigma} \to 0 \), learning becomes impossible and the expected time to approval becomes infinitely long. If \( \mathbb{P}(Z_\ell) > \mathbb{P}(Z_h) \), then \( R \)'s utility will converge to zero.

Next, we want to show that for \( c \) large enough, we can devise a (suboptimal) approval rule such \( \ell \) will drop out immediately and \( h \) will be approved with strictly positive probability. By the derivation of the optimal mechanism when \( \text{DIC}(Z_h, Z_\ell) \) is slack, we know that \( h \) will be approved with probability one. Moreover, as \( c \to \infty \), the function \( B_\ell(X) \to X \) and so the expected length of experimentation time goes to 0, giving \( R \) strictly positive utility.

\[ \square \]

E.7 Proof of Proposition 7

Proof. Suppose that \( A \) knows the state perfectly and \( R \) uses the symmetric mechanism for \( \pi = \mathbb{P}(Z_h) \). Then \( h \) will never have an incentive to quit early, since \( B \) is increasing in \( Z \). Moreover, by the same argument, \( \ell \) will choose to quit early. Therefore, let us define \( (\tau^h, d^h) \) to be the same as the symmetric mechanism and \( (\tau^\ell, d^\ell) \) to be the same as the symmetric mechanism except that it rejects whenever \( \ell \) would find it optimal to quit. This menu of mechanisms is clearly incentive compatible and using the \( (\tau^h, d^h) \) will yield the same distribution of approval times as the symmetric mechanism if \( \theta = H \) is present, but, \( (\tau^\ell, d^\ell) \) will lead to less approval than the symmetric mechanism if \( \theta = L \) and the same distribution over stopping times conditional upon approval. The value of this (suboptimal) asymmetric information mechanism is higher than that of the symmetric information case and so we can conclude that \( R \) strictly prefers \( A \) to learn the state.

\[ \square \]

F General Markov Process

We now move to the model of Section 6 and begin by proving a useful Lemma on the agent’s value function \( \hat{V} \).

Lemma 15. Under Assumption 2, for each \((b,B)\) such that \( B > X_{my}^A \), \( \hat{V}(B,b,X) \) satisfies strict single-crossing in \( X \) of \( f(X,0) \).
Proof. Since $B > X_{m_y}$ and $f(B, 1) > f(B, 0)$, we know that $\hat{V}(B, b, b) = f(b, 0)$ and $\hat{V}(B, b, B) > f(B, 0)$. For the sake of contradiction suppose that $\exists X$ such that for some $X' < X''$ such that $X \in (X', X'')$ and we had

$$\hat{V}(B, b, X') - f(X', 0) = \hat{V}(B, b, X'') - f(X'', 0) = 0 < \hat{V}(B, b, X) - f(X, 0).$$

We can rewrite $\hat{V}(B, b, X)$ as

$$\hat{V}(B, b, X) = \mathbb{E}[e^{-r\tau(X')}\hat{V}(B, b, X') 1(\tau(X') < \tau(X''))|X] + \mathbb{E}[e^{-r\tau(X'')}\hat{V}(B, b, X'') 1(\tau(X'') < \tau(X'))|X]$$

where the last line follows from the fact that pure delay is suboptimal. This contradicts the assumption that $\hat{V}(B, b, X) - f(X, 0) > 0$ and shows that we must have strict single crossing.

Before considering the problem $GSM$, we first consider the following constrained optimal stopping problem

$$\sup_{(\tau, d) \in \Delta_C} \mathbb{E}[e^{-r\tau}g(X_\tau, d)|X_0]$$

which is useful in the proof of Theorem 4 as well as Section 5.1) where $\Delta_C$ is the set of mechanisms such that

$$\forall X_i \in T_N \text{ such that } X_i \leq X_0 \text{ and } Y_m \in \mathcal{Y}_M$$

$$RDP(X_i) : \mathbb{E}[e^{-r(\tau(X_i))}f(X_{\tau(X_i)}, d_{\tau(X_i)})|X_0] \leq \mathbb{E}[e^{-r\tau}f(X_\tau, d_\tau)|X_0] \quad (5)$$

$$RPK(Y_m, X_i) : \mathbb{E}[e^{-r(\tau(X_i))}w_m(X_{\tau(X_i)}, d_{\tau(X_i)})|X_0] \leq Y_m.$$

where $T_N \in \mathbb{R}^{N+1}$ and $\mathcal{Y}_M \in \mathbb{R}^{M+1}$. This problem occupies the form of our relaxed problem where we drop all but a finite set of threshold quitting constraints. We can think of $g$ as the utility function for $R$ and $f$ as the utility function for $A$, while $w_m$ are a set of promise keeping constraints. We will assume that $g, f, w$ are all bounded in $X, d$ and that problem 5 satisfies all the conditions of Lemma 5. Let $H_N(X)$ be defined as the value of the optimal mechanism (in a relaxed problem) without $RPK$ constraints which makes sure that $A$ weakly wants to continue experimenting.
\[H_N(X) : \sup_{(\tau,d_\tau)} \mathbb{E}[e^{-r\tau} g(X_\tau, d_\tau)|X]\]

subject to \(\forall X_i \in T_N \cup \{X\} \) such that \(X_i \leq X\)

\[RDP(X_i) : \mathbb{E}[e^{-r(\tau \wedge \tau(X_i))} f(X_{\tau \wedge \tau(X_i)}, d_\tau(X_i))|X] \leq \mathbb{E}[e^{-r\tau} f(X_\tau, d_\tau)|X]\]

\[PK(X) : \mathbb{E}[e^{-r\tau} f(X_\tau, d_\tau)|X] \geq f(X, 0).\]

Let us define the set of \(RDP\) constraints which are binding when using the optimal \((\tau, d_\tau)\) for \(RSM_N\) as

\[B_N = \{X_i \in T_N : RDP(X_i) \text{ or, for some } m RPK(Y_m, X_i) \text{ is binding}\}.\]

We will assume write \(B_N = \{X^1, ..., X^{|B_N|}\}\) which are ordered from largest to smallest.

**Lemma 16.** Let \((\tau, d_\tau)\) be the solution to 5. Then the optimal stopping rule is a static threshold until \(X^1\) is reached for the first time; if there are no \(RPK\) constraints, then the continuation mechanism at \(\tau(X^1)\) solves \(H_N(X^1)\) where \(PK(X^1)\) binds.

**Proof.** By Lemma 5, there exists a set of multipliers \((\lambda(X_0), ..., \lambda(X_N))\) \(\in \mathbb{R}_+^{N+1}\) and \((\gamma(0, X_0), ..., \gamma(M, X_N))\) \(\in \mathbb{R}_+^{(M+1)(N+1)}\) such that the solution to \(\sup_{(\tau,d_\tau) \in \Delta_C} \mathbb{E}[e^{-r\tau} g(X_\tau, d_\tau)|X_0]\) is also a solution to

\[
\sup_{(\tau,d_\tau)} \mathbb{E}[e^{-r\tau} g(X_\tau, d_\tau)|X_0] \tag{6}
\]

\[- \sum_{j=1}^{|B_N|} \left[ \lambda(X^j) \left( \mathbb{E}[e^{-r(\tau \wedge \tau(X^j))} f(X_{\tau \wedge \tau(X^j)}, d_\tau(X^j))|X_0] - \mathbb{E}[e^{-r\tau} f(X_\tau, d_\tau)|X_0]\right) \right]
\]

\[+ \sum_{m=0}^M \gamma(m, X^j) \mathbb{E}[e^{-r(\tau \wedge \tau(X^j))} w_m(X_{\tau \wedge \tau(X^j)}, d_\tau(X^j))|X_0)].\]

We will argue that as long as \(X_t\) has not reached \(X^1\), the optimal policy must be a threshold policy where the process stops if \(X_t \geq B^1\) for some \(B^1 \geq 0\). To see this, define the continuation value of the optimal stopping rule after crossing \(X^1\) as
\[ k^R(X^1) := \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-rt}g(X_\tau, d_\tau) + \lambda(X^1)f(X_\tau, d_\tau) \right]
+ \sum_{j=2}^{[B_N]} \lambda(X^j) \left( e^{-r(\tau \wedge \tau^j[h_{\tau}(X^1)])} f(X_{\tau \wedge \tau(X^j)}, d_\tau(X^j)) \right) - e^{-rt} f(X_\tau, d_\tau) \right] \\
+ \sum_{m=0}^{M} \gamma(m, X^j) e^{-r(\tau \wedge \tau^j[h_{\tau}(X^1)])} w_m(X_{\tau \wedge \tau(X^j)}, d_\tau(X^j)) [X^1] \\
+ \lambda(X^1)f(X^1, 0) + \sum_{m=1}^{M} \gamma(j, X^1) w_m(X^1, 0). \] (7)

Note that at \( X^1 \) if \( \sum_{j=1}^{M} \gamma(j, X^1) w_m(X^1, 0) < 0 \), then \( R \) receives a one-time loss of \( \sum_{j=1}^{M} \gamma(j, X^1) w(X^1, 0) \) at exactly \( \tau(X^1) \), which might make it optimal to stop immediately. Therefore, we allow the continuation value at \( \tau(X^1) \) to be equal to the maximum of the payoff of continuing or stopping. This value \( K^R(X^1) \) is given by

\[ K^R(X^1) := \max \{ k^R(X^1), \max_{d_\tau} g(X_\tau, d_\tau) + \sum_{j=1}^{M} \sum_{m=1}^{[B_N]} \gamma(m, X^j) w_m(X^1, d_\tau) \}. \] (8)

The solution to \( K^R \) is independent of the previous history up until \( X^1 \) is reached. Since \( K^R \) is finite, we can use Assumption 2 to apply Proposition 5.7 from Dayanik and Karatzas (2003) to conclude that an optimal stopping rule to the problem defined such that the game ends when \( X^1 \) is first reached, yielding the stopping value \( K^R \). By the principal of optimality, we know that at time \( \tau(X^1) \), the value of optimal mechanism will be equal to \( K^R(X^1) \). Thus, treating \( K^R(X^1) \) as the continuation value upon reaching \( X^1 \) for the first time and dropping the constants \( Y_m \) from the problem, we can rewrite equation 6 as
\[
\sup_{(\tau, d_\tau)} \mathbb{E}[\mathbb{I}(\tau \leq \tau(X^1))e^{-r\tau}g(X_\tau, d_\tau) + \sum_{j=1}^{[\frac{[\mathbb{S}]}{\mathbb{N}}]} \lambda(X^j)(e^{-r(\tau \land X^j)})f(X_{\tau \land X^j}, d_\tau(X^j)) - e^{-r\tau}f(X_\tau, d_\tau)]
\]
\[
+ \sum_{m=0}^{M} \gamma(m, X^j)e^{-r(\tau \land X^j)}w_m(X_{\tau \land X^j}, d_\tau(X^j)) + \mathbb{I}(\tau > \tau(X^1))e^{-r\tau(X^1)}K^R(X^1)|X^0]
\]
\[
= \sup_{(\tau, d_\tau)} \mathbb{E}[\mathbb{I}(\tau \leq \tau(X^1))e^{-r\tau}g(X_\tau, d_\tau) + \sum_{m=0}^{M} \gamma(m, X^j)e^{-r(\tau \land X^j)}w_m(X_{\tau \land X^j}, d_\tau(X^j))] + \mathbb{I}(\tau > \tau(X^1))e^{-r\tau(X^1)}K^R(X^1)|X^0].
\]

By applying Lemma 4, we can see that the optimal stopping policy takes a threshold form which stops when \(X_t \geq B^1\) until \(\tau(X^1)\). If the policy stops at \(\tau(X^1)\), then we are done.

Therefore, suppose that the mechanism doesn’t end at \(\tau(X^1)\) and that \(RDP(X^1)\) is binding. The optimal stopping rule from \(\tau(X^1)\) onward is that which solves \(K^R(X^1)\).

Since this continuation mechanism does not depend on the history of play up until \(\tau(X^1)\), it must be that the constraint expectations over \(f\) at \(\tau(X^1)\)

\[
\mathbb{E}[e^{-r(\tau - \tau(X^1))}f(X_\tau, d_\tau)|X^1, h_\tau(X^1)],
\]

are also independent of the history to \(\tau(X^1)\) (and similarly for the other \(f\) constraint expectations). Let \(\tau^1 := \tau[h_\tau(X^1)], d^1 := d[h_\tau(X^1)]\) and \(\tau^1(X) := \tau^1[h_\tau(X^1)](X), d^1(X) := d^1[h_\tau(X^1)](X)\) respectively be the continuation mechanism and continuation threshold quitting rule at \(\tau(X^1)\). Then we can define the continuation value for \(A\) at \(\tau(X^1)\) as

\[
K^A(X^1) := \mathbb{E}[e^{-r\tau^1}f(X_{\tau^1}, d^1)|X^1].
\]

We can further decompose the initial expectation for \(f\) into what happens before and after \(\tau(X^1)\):

\[
\mathbb{E}[e^{-r\tau}f(X_\tau, d_\tau)|X^1] = \mathbb{E}[\mathbb{I}(\tau \leq \tau(X^1))e^{-r\tau}f(X_\tau, d_\tau)|X^0] + \mathbb{E}[\mathbb{I}(\tau > \tau(X^1))e^{-r\tau(X^1)}K^A(X^1)|X^0],
\]

(9)

and similarly for any \(X_i \leq X^1\)

\[
\mathbb{E}[e^{-r(\tau \land \tau(X^1))}f(X_{\tau \land \tau(X^1)}, d_\tau(X_i))|X^0] = \mathbb{E}[\mathbb{I}(\tau \leq \tau(X^1))e^{-r\tau}f(X_\tau, d_\tau)|X^0] + \mathbb{E}[\mathbb{I}(\tau > \tau(X^1))e^{-r\tau(X^1)}\mathbb{E}[e^{-r(\tau^1 \land \tau(X^1))}f(X_{\tau^1 \land \tau^1(X^1)}, d^1(X_i))|X^1]|X^0].
\]

(10)
Then we know by complementary slackness we know that for all $X^j$

$$
E[e^{-rt}f(X_t, d_t)|X_0] = E[e^{-r(\tau \wedge \tau(X^j))}f(\tau(X^j), d_\tau(X^j))|X_0].
$$

Substituting in equations 9, 10, we get

$$
E[\mathbb{1}(\tau > \tau(X^1))e^{-rt(X^1)}K^A(X^1)|X_0] = E[\mathbb{1}(\tau > \tau(X^1))e^{-r(\tau \wedge \tau(X^j))}f(\tau(X^1), d_{\tau(X^1)}(X^j)|X_0].
$$

Then, using the fact that neither $K^A(X^1)$ nor

$$
E[e^{-r(\tau \wedge \tau(X^j))}f(\tau(X^1), d_{\tau(X^1)}(X^j)|X_0]
$$

depend on the history up until $\tau(X^1)$ or the specific time of $\tau(X^1)$, we see that

$$
K^A(X^1) = E[\mathbb{1}(\tau > \tau(X^1))e^{-rt(X^1)}|X_0] = E[\mathbb{1}(\tau > \tau(X^1))e^{-r(\tau \wedge \tau(X^j))}f(\tau(X^1), d_{\tau(X^1)}(X^j)|X_0].
$$

Evaluating this at $X^j = X^1$, we see that

$$
K^A(X^1) = f(X^1, 0).
$$

Thus, upon reaching $X^1$, the expected continuation value to $A$ is equal to the value of quitting at $X^1$.

We now seek to show that the mechanism which solves $K^R(X^1)$ also solves $H_N(X^1)$ when there are no $RPK$ constraints. Let $(\tau^H, d^H_\tau)$ be the mechanism which solves $H_N$. Since $(\tau^1, d^1_\tau)$ satisfies all $RDP$ and $PK(0)$ constraints, it is admissible with respect to $H_N$. Therefore, it must be that

$$
E[e^{-rt^H}g(X_{\tau^H}, d^H_\tau)|X^1] \geq E[e^{-r\tau^1}g(X_{\tau^1}, d^1_\tau)|X^1]
$$

Moreover, since $(\tau^H, d^H_\tau)$ satisfies all $RPD$ constraints, switching from $(\tau^1, d^1_\tau)$ to $(\tau^H, d^H_\tau)$ at $\tau(X^1)$ is admissible in $\Delta_C$. Therefore we know that

67
sup \tau \text{ with approval at } B

\text{constraints are binding, then the optimal solution must be a static threshold mechanism.}

**Lemma 17.** When there are no RPK constraints, the solution to \( \sup_{(\tau, d_r) \in \Delta_C} \mathbb{E}[e^{-r\tau} g(X_\tau, d_r)] | X_0 \) is a static approval threshold policy until \( \tau(X^1) \) and doesn’t approve at \( \tau(X^1) \).

**Proof.** Let us suppose that at least one RDP constraint is binding (if no RDP constraints are binding, then the optimal solution must be a static threshold mechanism with approval at \( B \) and rejection at \( b \)). By Lemma 16, the optimal stopping rule until time \( \tau(X^1) \) is given by a static stopping rule \( \tau = \inf \{ t : X_t \geq B \} \) for some \( B \geq X_0 \). All that is left to derive is the decision rule.

First, let us suppose that \( R \) does not stop immediately at \( \tau(X^1) \) and, for the sake of contradiction, that \( R \) rejected at \( B \). By Lemma 16, the continuation value to \( A \) at \( \tau(X^1) \) is equal to his outside option and thus \( A \) receives the same utility as if \( R \) rejected at \( X^1 \): \( V(\tau, d_r, X_0) = V(\tau \geq (B) \wedge \tau(X^1), 0, X_0) \). But this will violate RDP\((X_0)\) since by Assumption 2

\[ \mathbb{E}[e^{-r(\tau \geq (B) \wedge \tau(b))} f(X_{\tau \geq (B) \wedge \tau(b)}, 0) | X_0] < f(X_0, 0). \]

Therefore, it must be that \( R \) is approving at \( \tau \geq (B) \) if \( \tau \geq (B) < \tau(X^1) \).

Now we consider the case in which \( R \) also stops at \( X^1 < X_0 \) (the constraint at \( X^1 \) will be binding if \( R \) acts at \( X^1 \)). We can construct the Lagrangian corresponding to GRSM\(_N\) as

\[ \mathcal{L} = \sup_{(\tau, d_r)} \mathbb{E}[e^{-r\tau} g(X_\tau, d_r)] + \sum_{j=1}^{|B_N|} \lambda(X^j) (e^{-r(\tau \wedge \tau(X^j))} f(X_{\tau \wedge \tau(X^j)}, d_r(X^j)) - e^{-r} f(X_\tau, d_r)) | X_0]. \]
Since we can now analyze the problem as a single-decision maker would, it is easy to see that

\[ d_\tau(X_1) = 1 \iff g(X_1, 1) > g(X_0, 0) \Rightarrow g(B, 1) > g(B, 0) \iff d_{\tau \geq B} = 1. \]

Therefore, if the optimal decision rule approves at \( X_1 \), then it must approve at both \( B \) as well. But by Assumption 2, we know that immediate approval would be better for \( R \) than waiting to approve. Since this is also better for \( A \), immediate approval would be admissible with respect to \( GRSM_N \), contradicting the optimality of waiting until \( B \) or \( X_1 \).

Lemma 18. Under Assumptions 2 and 3, as \( N \to \infty \), the solution to \( H_N(X^1) \) (when \( PK(0) \) is binding) is given by \( \tau = \inf\{t : X_t \geq \underline{B}(M_t^X)\} \land \tau(b_r) \) and \( d_\tau = \mathbb{1}(X_t = \underline{B}(M_t^X)) \) for some cutoff \( b_r \).

Proof. We now want to solve for the optimal mechanism which solves \( H_N(X) \) when \( PK(X) \) is binding. By applying Lemma 16 repeatedly, we see that the optimal mechanism will consist of a sequence of upper stopping thresholds \( \{\underline{B}_n\}_{n=1}^{B_N} \) and one lower threshold \( b_r \). Moreover, at each binding constraint \( X^j > X^{[B_N]} \), we must have \( PK \) binding at \( X^j \). As the mechanism progresses, the current upper threshold depends only on the lowest binding threshold which has been reached; hence, we can write the upper threshold at time \( t \) as a step function \( B_N(M_t^X) \) of the minimum of the \( X \) until time \( t \).

Suppose that the state is \( X_t = M_t^X = X^j \) and \( X^j > X^{[B_N]} \). By complementary slackness, we know that at \( X^j \) with continuation mechanism \( (\tau^j, d^j_\tau) \) we have

\[
\mathbb{E}[e^{-r\tau^j} f(X_{\tau^j}, d^j_\tau)|X^j] = f(X^j, 0)
\]

Therefore, in order to satisfy \( RDP \) at \( X^{j+1} \), we must have approval at \( B_N(X^{j+1}) \), since, if \( R \) rejected at this threshold then

\[
\mathbb{E}[e^{-r\tau} f(X_\tau, d_\tau)|X^j] = \mathbb{E}[\mathbb{1}(\tau \leq \tau(X^{j+1})) e^{-r\tau} f(X_\tau, d_\tau)|X^j] \\
+ \mathbb{E}[\mathbb{1}(\tau > \tau(X^{j+1})) e^{-r\tau(X^{j+1})}\mathbb{E}[e^{-r(\tau - \tau(X^{j+1}))} f(X_\tau, d_\tau)|X^{j+1}]|X^j] \\
= \mathbb{E}[e^{-r(\tau(B_i) \land \tau(X^{j+1}))} f(X_{\tau(B_i) \land \tau(X^{j+1})}, 0)|X^j] \\
< f(X_i, 0).
\]
a violation of RDP at $X^j$. Therefore, since $R$ is approving at $B$ and $PK$ is binding, the threshold $B$ at $X^j$ must be such that $\hat{V}(B, X^{j+1}, X^j) = 0$.

We now turn to the relationship between $X^j$ and $X^{j+1}$. Intuitively, we should not expect them to be far apart: Between time $\tau(X^j)$ and $\tau(X^{j+1})$, the optimal mechanism yields $A$’s outside option to $A$ whenever $X_t = X^j$ (since, for $A$, the mechanism is equivalent to a static threshold with approval at $B$ and rejection at $b$) and when $X_t = X^{j+1}$ (by complementary slackness). Therefore, when $X_t \in (X^{j+1}, X^j)$, the mechanism is, for $A$, equivalent to a static threshold mechanism (with thresholds $(X^{j+1}, X^j)$) giving him his outside option at both thresholds. Therefore, he would prefer taking his outside option at such $X_t$. Thus, if there is a constraint $X_i \in T_N$ such that $X_i \in (X^{j+1}, X^j)$, this will violate the RDP constraints for large $N$ by the suboptimality of delay and the fact eventually some $X^i$ RDP constraint will fall in $(X^j, X^{j+1})$.

We can also show that at the final binding quitting threshold $X^{[B_N]}$, $R$ must be rejecting. After time $\tau(X^{[B_N]})$, there are no more binding constraints and the optimal stopping mechanism will solve

$$
\sup_{(\tau, d_t)} \mathbb{E}[e^{-\tau \tau} (g(X_\tau, d_\tau) - \sum_{j=1}^{[B_N]} \lambda(X^j) f(X_\tau, d_\tau))) X^{[B_N]}].
$$

By Lemma 4, the solution to this problem will be a pair of threshold $(b, B)$ and, using arguments from Lemma 17, it must be that $R$ approves at $B$ and rejects at $b$ (which implies that $X^i < b$ are binding).

These results tell us that for each $N$, the optimal approval rule is a function $B_N(M^X_t)$ which is constant when $M^X_t \in [X^{j+1}, X^j)$. Because $\hat{V}$ is single-peaked, there are at most two $B$ which satisfy $\hat{V}(B, X^{j+1}, X^j) = f(X^j, 0)$. If there do exist two thresholds, then it must be that one threshold (call it $B_1$) is above $B^*_A(X^{j+1})$ and one (call it $B_2$) is below $B^*_A(X^{j+1})$. Taking $N$ large, we argue that the optimal threshold must always be $B_1$.

First, we argue that when $M^X_t \in [X^{j+1}, X^j]$, it must be that $R$’s preferred approval threshold is higher than that of $A$. By Assumption 4, we know that if $R$ were to expect rejection at $\tau(X^{j+1})$ his preferred approval threshold (call it $\hat{B}$) would be higher than that of $A$. Under the optimal mechanism, when she has a higher utility of $J(X^{j+1}) > g(X^{j+1}, 0)$ at $\tau(X^{j+1})$, which strictly increase $R$’s value function and, we argue, increases his preferred approval threshold (call it $B^*$). For the sake of contradiction, suppose that $B^* < \hat{B}$. This is implies that $R$’s value function at $B^*$ when she expects rejection at $\tau(X^{j+1})$ is greater than that of approving immediately, which is what her value function is when she expects to have continuation value $J(X^{j+1})$, contradicting the fact the fact that her value function strictly increased. Therefore, it must be that $B^* > \hat{B}$.
Now suppose that \( R \) were to use the threshold \( B_2 \). Then she could increase the threshold to \( A \)'s preferred threshold when \( A \) expects to reject at \( \tau(X^{j+1}) \), which would weaken DP constraints as well as increase \( R \)'s utility since it moves the approval threshold closer to \( R \)'s preferred point. Therefore it must be that \( B = B_1 \).

We now turn to examine \( \lim_{N \to \infty} B_N(M^X_t) \). Since \( X^{[B_N]} \in [X, X^1] \), there exists a limit of \( X^{[B_N]} \) as \( N \to \infty \). Let \( \overline{X}(M^X_t) := \max\{X_i \in T_N : X_i < M^X_t\} \). By the previous arguments we have \( B_N(M^X_t) = \overline{B}(\overline{X}(M^X_t), X(M^X_t) - \frac{1}{N}) \). Since \( \overline{B}(X_j, X_i) \) is continuously differentiable, \( B_N(M^X_t) \) has uniformly bounded variation on \([X, X^1]\) and is uniformly bounded, by Helly’s Selection Theorem, it has a limit. Since RDP binding at \( X_t \) implies that it is binding at \( X_{t+1} = X_t - \frac{1}{N} \), we can see that on \( M^X_t > X^{[B_N]} \) we have

\[
\lim_{N \to \infty} B_N(M^X_t) = \overline{B}(\overline{X}(M^X_t), X(M^X_t) - \frac{1}{N}) = B(M^X_t)
\]

\[\square\]

**Lemma 19.** Let \((\tau^N, d^N_t)\) be the solution to \( GRSM_N \) and define \((\tau, d_r) = \lim_{N \to \infty} (\tau^N, d^N_t)\). Then \((\tau, d_r)\) is admissible with respect to \( GSM \).

**Proof.** Our stopping rule \( \tau = \inf \{t : X_t \geq B(M^X_t)\} \land \tau(b_r) \) is clearly a \( F^X_t \)-measurable and thus is admissible if it satisfies all DP constraints. We need to verify that after any history, the continuation value for \( A \) is weakly positive. Since the mechanism \((\tau, d_r)\) is Markov with respect to \((X_t, M^X_t)\), we need only check that \( \mathbb{E}[e^{-r \tau} f(X_t, d_r)|X_t, M^X_t] \geq f(X_t, 0) \).

When \( M^X_t > b^*(B^1) \), then \( A \)'s continuation utility is given by \( \tilde{V}(B^1, b^*(B^1), X_t) \), which is greater than \( f(X_t, 0) \) by definition of \( b^*(B^1) \) and the fact that \( X_t > b^*(B^1) \). When \( M^X_t \leq b^*(B^1) \), then \( A \)'s continuation utility is given by \( \tilde{V}(B(M^X_t), M^X_t, X_t) \), which is greater than \( f(X_t, 0) \) since \( X_t \geq M^X_t \). Thus, for all \((X_t, M^X_t)\), dynamic participation constraints are satisfied.

\[\square\]

**F.1 Proof of Theorem 4**

**Proof.** Let \((\tau, d_r) := \lim_{N \to \infty} (\tau^N, d^N_t)\) where \((\tau^N, d^N_t)\) solves \( GRSM_N \). We know that the value of \( GRSM_N \) (i.e., \( J(\tau^N, d^N_t, X_0) \)) is an upper bound on the value of \( GSM \). Moreover, since the \( J(\tau^N, d^N_t, X_0) \) is bounded above by \( \sup_{(\tau, d_r)} \mathbb{E}[e^{-r \tau} f(X_t, d_r)|X_0] \), the dominated convergence theorem implies that \( J(\tau, d_r, X_0) = \lim_{N \to \infty} J(\tau^N, d^N_t, X_0) \). Therefore, we can conclude that \( J(\tau, d_r, X_0) \) is an upper bound on the value of \( GSM \). Since
$(\tau, d_\tau)$ is admissible to $GSM$ by Lemma 19, this implies that $(\tau, d_\tau)$ is a solution to $GSM$.

Finally, we need to show that the threshold $B(M_1^X)$ is continuous. By Lemma 16, we know that the approval threshold (call it $B^1$) is constant until $\tau(X^1)$. After $\tau(X^1)$, Lemma 11 implies that $B(M_1^X) = B(M_1^X)$. Thus, we will be done if we can show that $X^1 = b^*(B^1)$.

Clearly, we cannot have $X^1 < b^*(B^1)$, since this would violate the RDP constraints for $X_i \in (X^1, b^*(B^1))$. For the sake of contradiction, suppose that for large enough $N$ we had $X^1 > b^*(B^1)$ where $B^1$ is the initial threshold for $GRSM_N$. Let $\hat{J}(X_c)$ be the continuation value of the optimal mechanism $(\tau^N, d_c^N)$ of $GRSM_N$ to $R$ at some $X_c \in (X^1, X_0)$ when $M_1^X > X^1$ and let $\tilde{J}(X_c)$ be the continuation value to $R$ at $X_c$ when $M_1^X \in (X^2, X^1)$. The utility to $R$ at $t = 0$ is given by

$$
E[\mathbb{1}(\tau \leq \tau(X_c))e^{-\mathbb{1}(\tau(X_c))\mathbb{1}(\tau(X_c))\hat{J}(X_c)|X_0, M_1^X = X_0}].
$$

Switching to the mechanism which delivers $\tilde{J}(X_c)$ at $\tau(X_c)$ would be admissible since $\hat{J}(X_c)$ satisfies all RDP constraints. Therefore, for this to not be optimal, it must be that $\tilde{J}(X_c) < \hat{J}(X_c)$. Similarly, consider the continuation value of the optimal mechanism at $\tau(X^1)$, which is given by

$$
E[\mathbb{1}(\tau(X^2) \leq \tau(X_c))e^{-\mathbb{1}(\tau(X^2))\mathbb{1}(\tau(X^2))\tilde{J}(X_c)|X^1, M_1^X = X^1}].
$$

We can see that switching to the mechanism which delivers $\tilde{J}(X_c)$ if $X_c$ is reached before $\tau(X^2)$ will satisfy all RDP at $\tau(X^1)$ (since the mechanism at $\tilde{J}(X_c)$ satisfies all RDP constraints). Therefore, for this to not be optimal, we must have $\tilde{J}(X_c) > \hat{J}(X_c)$, a contradiction. We can conclude that $X^1 = b^*(B^1)$.

\[ \square \]

### G Two-Sided Commitment

Under two-sided commitment, $R$ proposes a binding contract to $A$ which specifies the amount of experimentation that $A$ must perform. However, unlike the principal optimal solution, $A$ has some say in the design of the mechanism: More specifically, $A$ has authority to accept or reject the mechanism at $t = 0$ (and at that time alone). If $A$ accepts the mechanism, then $A$ commits to continue experimentation until the mechanism specifies that experimentation ends. We define this environment below:
Definition 11. A mechanism has two-sided commitment if once A has accepted $(\tau, d_\tau)$, experimentation continues until $\tau$.

Since A has the option of rejecting R’s proposed mechanism, R’s mechanism must satisfy a participation constraint. Formally, this means that A’s expected utility from R’s proposed mechanism must be at least as high as A’s outside option, which we take to be 0, when rejecting R’s mechanism.

Definition 12. A mechanism $(\tau, d_\tau)$ satisfies the participation constraint if $V(\tau, d_\tau, Z_0) \geq 0$.

Since A has no private information about the state and the contract, once agreed to, is binding, the participation constraint will be the sole constraint on R’s choice of a mechanism. Formally, the mechanism design problem faced by R is given by

$$\sup_{(\tau, d_\tau)} \mathbb{E}[e^{-\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_0]$$

subject to

$$P : \mathbb{E}[e^{-\tau} (d_\tau \frac{e^{Z_\tau} + a + c}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_0] - \frac{c}{r} \geq 0,$$

where $P$ is a participation constraint that ensures that the agent finds it optimal to agree to the mechanism.

Our problem takes the form of a constrained optimal stopping problem. A robust finding from the single decision-maker problem (and as seen in our principal-optimal solution) is the optimality of static-threshold rules (e.g., Wald (1947) and Moscaroni and Smith (2001)). However, with agency considerations the participation constraint prevents the use of standard time-consistency arguments that imply the optimality of threshold rules. Despite this difficulty, we are able to show in Proposition 8 that static threshold mechanisms remain optimal.

Proposition 8. The solution to the symmetric information problem with two-sided commitment takes the form of a static-threshold policy. If $b \neq -\infty$, then the optimal approval and rejection thresholds $(B, b)$ are the solution to the following equations:

$$\Psi_b(1 - e^{-B}) = e^{-B}\Psi \left(\frac{e}{r}[\psi_b(1 + e^{-B}) - e^{-B}] + \Psi_b[(1 + \frac{c}{r} + (a + \frac{c}{r})e^{-B})] \right) = a\Psi e^{-B} + \Psi B(1 + \frac{c}{r} + (a + \frac{c}{r})e^{-B}) + \psi_B(1 + e^{-B})\frac{e}{r}$$

$$\Psi(1 + \frac{c}{r} + (a + \frac{c}{r})e^{-B}) + \frac{c}{r}\psi(1 + e^{-B}) = \frac{c}{r} \frac{1 + e^{Z_0}}{e^{Z_0}}$$

where $B < B^{FB}$ and $P(\tau, d_\tau)$ is binding. If $b = -\infty$, then $B = B^{FB}$.
Proof. We start by proving the conditions of Lemma 5 are met. To see this, note that
the stopping policy \( \tau = \epsilon \) and \( d_\tau = 1 \) will keep the participation constraint slack for \( \epsilon \)
small enough. The other conditions of Lemma 5 are easily checked.

By applying Lemma 5, we can use a Lagrangian in order to turn the primal problem:

\[
\sup_{(\tau,d_\tau)} \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_0]
\]

subject to

\[
P : \quad \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_0] - \frac{c}{r} \geq 0,
\]

into the dual problem

\[
\mathcal{L} = \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_0] + \lambda \left[ \frac{c}{r} - \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_0] \right]
\]

\[
= \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} - \lambda \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} - \lambda \frac{c}{r}) | Z_0] + \lambda \frac{c}{r}.
\]

By Lemma 4, we can verify that the solution is of a threshold form. Let \((B, b)\) be the
approval and rejection threshold respectively. Then we know that the primal problem
must solve

\[
\mathcal{L} = \Psi \left( \frac{e^{Z_0}(1 - e^{-B})}{1 + e^{Z_0}} \right) - \lambda \Psi \left( \frac{e^{Z_0}(1 + \frac{a + \frac{c}{r}e^{-B}}{1 + e^{Z_0}})}{1 + e^{Z_0}} \right) - \lambda \frac{c}{r} \frac{e^{Z_0}(1 + e^{-b})}{1 + e^{Z_0}}.
\]

Taking first-order conditions are rearranging yields the equality in the proposition. \(\Box\)

We see that, as long as the the principal optimal mechanism is not achievable, \( A \)
must be indifferent between accepting the mechanism and taking his outside option. Any
mechanism in which the rejection threshold is not \(-\infty\) and \( A \) strictly prefers to
accept the mechanism can be improved upon by lowering the rejection threshold slightly,
thereby increasing \( R \) utility (since experimentation is always valuable to \( R \)) while
preserving the participation constraint. This result also establishes that the solution under
two-sided commitment is qualitatively the same as in the single decision-maker (thereby
justifying some of the focus on threshold mechanisms in the literature), albeit quantita-
tively different. This quantitative difference is described in the following corollary.

Corollary 1. If \((B^TS, b^TS)\) is the optimal two-sided commitment thresholds, then \( B^{FB} \geq B^TS \) and \( b^{TS} \geq b^{FB} \).
Whenever there is rejection in the two-sided commitment problem, the option value of experimentation is lower in the two-sided commitment case than in the principal-optimal case. Since with two-sided commitment experimentation will be ended earlier at low beliefs relative to the principal-optimal, the value of experimentation at intermediate beliefs decreases relative to approval leading to a lower approval threshold.

Interestingly, as in Henry and Ottaviani (2018), the choice of \( B, b \) will depend on the initial \( Z_0 \), which differs from the single-decision maker problem. The introduction of the participation constraint at \( t = 0 \) means that the pure time-consistency of single-decision maker problems is no longer present.

### H Quantitative Derivation

Define the function \( j_i(X, M, b_r) \) (we will drop \( b_r \) for notational convenience) to be the expected value of the principal when the current minimum of evidence is \( M \), current beliefs are \( Z = Z_i + \frac{\phi}{\sigma} X \) and the project is rejected when beliefs reach \( b_r \). Using our previous formulas for discounted threshold crossing probabilities, it is easy to see that

\[
j_i(X, M^X) = \Psi(B_i, Z(M^Z), M^Z, Z) \frac{e^Z - e^{Z-B_i, Z(M^Z)}}{1 + e^Z} \]

Expected utility from approval before \( Z_t = M^Z \)

\[
+ \psi(B_i, Z(M^Z), M^Z, Z) \frac{e^Z + e^{Z-M^Z}}{1 + e^Z} \cdot j_i(M^X, M^X) \]

Continuation value at \( M^X \)

where the mapping from \( X, M^X, B \) to \( X, M^Z, B \) is understood. Thus if we can calculate \( j_i(M^X) := j_i(M^X, M^X) \), the value of \( j_i(X, M^X) \) follows immediately. In order to calculate \( j_i(M^X) \), we use the principle of normal reflection\(^{16}\): \( \frac{\partial j_i(X, M^X)}{\partial M^X}|_{X=M^X} = 0 \). We can then take the derivative with respect to \( M^X \) to get

\[
\frac{\partial j_i(X, M^X)}{\partial M^X} = \frac{\phi}{\sigma} B_i, Z(M^Z) \left[ \Psi_B \frac{e^Z - e^{Z-B_i, Z(M^Z)}}{1 + e^Z} + \Psi \frac{e^{Z-B_i, Z(M^Z)}}{1 + e^Z} + \psi_B \frac{e^Z + e^{Z-M^Z}}{1 + e^Z} \right] j_i(M^X) \\
+ \frac{\phi}{\sigma} \psi_b \frac{e^Z - e^{Z-B_i, Z(M^Z)}}{1 + e^Z} + \psi_b \frac{e^Z + e^{Z-b}}{1 + e^Z} j_i(M^X) - \psi \frac{e^{Z-M^Z}}{1 + e^Z} j_i(M^X) \\
+ j_i'(M^X) \psi \frac{e^Z + e^{Z-M^Z}}{1 + e^Z}.
\]

\(^{16}\)See Peskir and Sharyaev (2006) for a derivation.
Evaluating the above equation at $Z = M^Z$ and using that $\frac{\partial j_i(X,M^X)}{\partial M^X}|_{Z=M^+} = 0$, we get

$$j'_i(M^X) = j_i(M^X)\left[\frac{1}{1+e^{M^X}} - \psi_b\right] - \frac{e^{M^X} - e^{M^X-B_{i,Z}(M^Z)}}{1+e^M} \Psi_b,$$

(11)

where we note that $\Psi(B(M), M, M) = 0$ and $\psi(B(M), M, M) = 1$. This, coupled with the boundary condition $j_i(b_r) = 0$ gives the ODE which describes $j_i(M)$.

**Proposition 9.** The value of experimentation to $R$ for type $i$ when the current evidence level is $M^X_t$ and the minimum is $M^X_t$ is given the unique solution to $j'_i(M^X_t)$.

**Proof.** Let $j_i(X, M^X)$ be given by

$$j_i(X, M^X) = \Psi(B_{i,Z}(M), M, Z_i) \frac{e^{Z_i} - e^{Z_i-B_{i,Z}(M)}}{1+e^{Z_i}} + \psi(B_{i,Z}(M), M, Z_i) \frac{e^{Z_i} + e^{Z_i-M}}{1+e^{Z_i}} j_i(M),$$

which is the solution to the Dirichlet problem (dropping $i$ subscripts)

$$\mathbb{L}_X j_i(X, M^X) = r j_i(X, M),$$

$$j_i(B(M), M) = B(M),$$

$$j_i(M, M) = j_i(M),$$

and $j_i(M)$ is the solution to the differential equation (derived using the principle of normal reflection $\frac{\partial j_i(X,M^X)}{\partial M^X}|_{X=M^X = 0}$)

$$j'_i(M^X) = \frac{\phi}{\sigma} \left( j_i(M^X)\left[\frac{1}{1+e^{M^X}} - \psi_b B'_Z(M) - \psi_b\right] - \frac{e^{M^Z} - e^{M^Z-B'_Z(M^Z)}}{1+e^{M^Z}} [\Psi B'_Z(M^Z) + \Psi_b] \right),$$

(12)

with boundary condition $j_i(M) = 0$.

We first argue that there is a unique solution to the differential equation for $j'_i(M)$. Lipschitz continuity of the RHS of equation 12 is clear and continuity in $M$ follows from the continuity of $B(M), B'_Z(M)$. Therefore, the Picard-Lindelof Theorem implies a unique solution.

Now we want to argue that $j_i(X_0, M^X_0) = \mathbb{E}[e^{-\tau r} d_t e^{B_Z(M^Z)/1+e^{B_Z(M^Z)}} | Z_0, M_0]$ where $\tau = \inf\{t: X_t \geq B(M^X_t)\} \wedge (\phi/\sigma) Z_i$. By applying Ito’s Lemma to $j_i(X, M^X)$, we have...
\[ e^{-rt} j(X_t, M_t^X) = j(Z_0, M_0^X) + \int_0^t e^{-rs} \left[ \sigma \frac{\partial j(X_s, M_s^X)}{\partial X} dB_s + \frac{\partial^2 j(X_s, M_s^X)}{\partial X^2} \frac{\sigma^2}{2} ds - r j(X_s, M_s^X) ds + \frac{\partial j(X_s, M_s^X)}{\partial M_s^X} dM_s^X \right] \]

\[ = j(X_0, M_0^X) + S_t + \int_0^t e^{-rt} [L_X j(X_s, M_s^X) - r j(X_s, M_s^X)] ds, \]

where we use the fact that \( \frac{\partial j(M_s^X, M_s^X)}{\partial M_s^X} = 0 \) and \( \Delta M_s^X = 0 \) when \( X_s > M_s^X \) and we define \( S_t \) to be

\[ S_t = \int_0^t e^{-rt} \frac{\partial j(X_s, M_s^X)}{\partial X} dB_s, \]

which is a continuous local martingale.

We now note that \( L_X j(Z_s, M_s^X) - r j(X_s, M_s^X) = 0 \) for all \( X_s \in (M_s^X, B(M_s^X)) \).

Therefore, we can reduce the above equation for \( e^{-rt} j(X_t, M_t^X) \) to

\[ e^{-rt} j(X_\tau, M_\tau^X) = j(X_0, M_0^X) + S_\tau. \]

When the process is stopped, the value \( j(X_\tau, M_\tau^X) \) is always equal to \( \mathbb{1}(X_\tau \geq B(M_\tau^X)) \frac{e^{B_2(M_\tau^X)} - 1}{1 + e^{B_2(M_\tau^X)}} \). Therefore, we have that

\[ e^{-rt} \mathbb{1}(X_\tau \geq B(X_t^X)) \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} = e^{-rt} j(X_\tau, M_\tau^X) = j(X_0, M_0^X) + S_\tau. \]

Taking expectations of both sides, we have

\[ \mathbb{E}[e^{-rt} \mathbb{1}(X_\tau = B(M_t^X)) \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_0] = j(X_0, M_0^X) + \mathbb{E}[S_t | Z_0]. \]

It follows from Doob’s optimal sampling theorem that \( \mathbb{E}[S_t | Z_0, M_0^X] = 0. \) Noting that \( \mathbb{E}[e^{-rt} d_T \frac{e^{B_2(M_t^X)} - 1}{1 + e^{B_2(M_t^X)}} | Z_0, M_0^X] = \mathbb{E}[e^{-rt} \mathbb{1}(X_\tau \geq B(M_t^X)) \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_0] \), we can conclude that \( j(X_0, M_0^X) = \mathbb{E}[e^{-rt} d_T \frac{e^{B_2(M_t^X)} - 1}{1 + e^{B_2(M_t^X)}} | Z_0, M_0^X] \):

By a similar strategy, we can solve for \( \ell, h \)'s value function using when their incentive constraints begin to bind as a boundary condition. Once we have solved for the relevant differential equations, we need only optimize over at most six parameters.